

Complex Land Cover Classifications and Physical Properties Retrieval of Tropical Forests using Multi-Source Remote Sensing

By the Faculty of Geosciences, Geo-Engineering and Mining
of the Technische Universität Bergakademie Freiberg

approved

THESIS

to attain the academic degree of

Doctor rerum naturalium

(Dr.rer.nat.)

Submitted

by **MSc Arief Wijaya**

born on the 24 March 1975 in Jakarta, Indonesia

Assessors: Prof. Dr. Lothar Ratschbacher, Freiberg, Germany
Prof. Dr. Hermann Heilmeier, Freiberg, Germany
Prof. Dr. Totok Gunawan, Yogyakarta, Indonesia
Dr. Richard Gloaguen, Freiberg, Germany
Dr. Irmgard Niemeyer, Jülich, Germany

Date of the award: 30 April 2010

Abstract

The work presented in this thesis mainly focuses on two subjects related to the application of remote sensing data: (1) for land cover classification combining optical sensor, texture features generated from spectral information and synthetic aperture radar (SAR) features, and (2) to develop a non-destructive approach for above ground biomass (AGB) and forest attributes estimation employing multi-source remote sensing data (i.e. optical data, SAR backscatter) combined with in-situ data. Information provided by reliable land cover map is useful for management of forest resources to support sustainable forest management, whereas the generation of the non-destructive approach to model forest biophysical properties (e.g. AGB and stem volume) is required to assess the forest resources more efficiently and cost-effective, and coupled with remote sensing data the model can be applied over large forest areas. This work considers study sites over tropical rain forest landscape in Indonesia characterized by different successional stages and complex vegetation structure including tropical peatland forests. The thesis begins with a brief introduction and the state of the art explaining recent trends on monitoring and modeling of forest resources using remote sensing data and approach. The research works on the integration of spectral information and texture features for forest cover mapping is presented subsequently, followed by development of a non-destructive approach for AGB and forest parameters predictions and modeling. Ultimately, this work evaluates the potential of mosaic SAR data for AGB modeling and the fusion of optical and SAR data for peatlands discrimination. The results show that the inclusion of geostatistics texture features improved the classification accuracy of optical Landsat ETM data. Moreover, the fusion of SAR and optical data enhanced the peatlands discrimination over tropical peat swamp forest. For forest stand parameters modeling, neural networks method resulted in lower error estimate than standard multi-linear regression technique, and the combination of non-destructive measurement (i.e. stem number) and remote sensing data improved the model accuracy. The up scaling of stem volume and biomass estimates using Kriging method and bi-temporal ETM image also provide favorable estimate results upon comparison with the land cover map.

Zusammenfassung

Die in dieser Dissertation präsentierten Ergebnisse konzentrieren sich hauptsächlich auf zwei Themen mit Bezug zur angewandten Fernerkundung: 1) Der Klassifizierung von Oberflächenbedeckung basierend auf der Verknüpfung von optischen Sensoren, Textureigenschaften erzeugt durch Spektraldaten und Synthetic-Aperture-Radar (SAR) features und 2) die Entwicklung eines nichtdestruktiven Verfahrens zur Bestimmung oberirdischer Biomasse (AGB) und weiterer Waldeigenschaften mittels multi-source Fernerkundungsdaten (optische Daten, SAR Rückstreuung) sowie in-situ Daten. Eine zuverlässige Karte der Landbedeckung dient der Unterstützung von nachhaltigem Waldmanagement, während eine nichtdestruktive Herangehensweise zur Modellierung von biophysikalischen Waldeigenschaften (z.B. AGB und Stammvolumen) für eine effiziente und kostengünstige Beurteilung der Waldressourcen notwendig ist. Durch die Kopplung mit Fernerkundungsdaten kann das Modell auf große Waldflächen übertragen werden. Die vorliegende Arbeit berücksichtigt Untersuchungsgebiete im tropischen Regenwald Indonesiens, welche durch verschiedene Regenerations- und Sukzessionsstadien sowie komplexe Vegetationsstrukturen, inklusive tropischer Torfwälder, gekennzeichnet sind. Am Anfang der Arbeit werden in einer kurzen Einleitung der Stand der Forschung und die neuesten Forschungstrends in der Überwachung und Modellierung von Waldressourcen mithilfe von Fernerkundungsdaten dargestellt. Anschließend werden die Forschungsergebnisse der Kombination von Spektraleigenschaften und Textureigenschaften zur Waldbedeckungskartierung erläutert. Desweiteren folgen Ergebnisse zur Entwicklung eines nichtdestruktiven Ansatzes zur Vorhersage und Modellierung von AGB und Waldeigenschaften, zur Auswertung von Mosaik- SAR Daten für die Modellierung von AGB, sowie zur Fusion optischer mit SAR Daten für die Identifizierung von Torfwäldern. Die Ergebnisse zeigen, dass die Einbeziehung von geostatistischen Textureigenschaften die Genauigkeit der Klassifikation von optischen Landsat ETM Daten gesteigert hat. Desweiteren führte die Fusion von SAR und optischen Daten zu einer Verbesserung der Unterscheidung zwischen Torfwäldern und tropischen Sumpfwäldern. Bei der Modellierung der Waldparameter führte die Neural-Network-Methode zu niedrigeren Fehlerschätzungen als die multiple Regressions. Die Kombination von nichtdestruktiven Messungen (z.B. Stammzahl) und Fernerkundungsdaten führte zu einer Steigerung der Modellgenauigkeit. Die Hochskalierung des Stammvolumens und Schätzungen der Biomasse mithilfe von Kriging und bi-temporalen ETM Daten lieferten positive Schätzergebnisse im Vergleich zur Landbedeckungskarte.

Acknowledgments

I think this is the right moment to acknowledge everybody who supported me to complete this doctoral thesis. Firstly, I would like to thank my supervisor, Dr. Richard Gloaguen, the Head of Remote Sensing Group, TU-Bergakademie Freiberg, for his tremendous supports, both academically and personally. His great contributions on this work and number of great discussions, as well as the way he motivates me to actively participate in international conferences and his critical comments on my scientific publications, I have to acknowledge as well.

I would like to acknowledge Prof. Dr. Hermann Heilmeyer, my second supervisor from Interdisciplinary Ecological Centre, TU-Bergakademie Freiberg, who provides me wonderful supports and time for discussion and for the manuscript corrections.

Deutscher Akademischer Austauschdienst (DAAD) deserves special mention for giving me the scholarship to conduct doctoral study at TU-Bergakademie Freiberg. Personally, I would like to thank my DAAD referat, Mrs. Susanne Kammüller, Mrs. Helga Islam and Mrs. Hannelore Bossmann for their valuable supports during my stay in Germany.

Dr. Prashanth Reddy Marpu, who helped me in many aspects included in this thesis also for our friendship, thank you. For my colleagues in Remote Sensing Group: Veraldo Liesenberg, Syed Amer Mahmood, Faisal Shahzad, Moncef Bouaziz, and Mathias Leidig, thank you for walking along with me during my study in Freiberg and for many great discussions, as well as for an enjoying moment in Cape Town, I will never forget it. Special thanks to Anja Bretzler for English corrections and abstract translation. For my Indonesian friends: Pak Eka, Pak Gito, Sandi, Alin, Widia, Yesi dan Budi, I would also like to thank for the friendship.

To my both parents, H. Harsono and Hj. Wuryaningsih who always pray for me to achieve this level, for their great supports all my life, also for their love and care. To my wife, Ratna, and my lovely daughter, Keisya, thank you for sharing the sadness and happiness with me in Freiberg, also for all your love and pray, and for our precious moment together.

Table of Contents

Abstract	i
Zusammenfassung	ii
Acknowledgments	iii
Table of Contents	iv
List of Figures	ix
List of Abbreviations	xiv
Chapter 1 Introduction	1
1.1 Tropical forest disturbance and monitoring	2
1.2 Deforestation in Indonesia and its Impacts on Carbon Storage	2
1.3 Research Questions	4
1.4 Scientific Challenge and Objectives	4
1.5 Materials and Method	5
1.6 The Study Areas	7
1.6.1 Labanan concession forest (study site 1)	7
1.6.2 Tanjung Puting national park (study site 2)	11
1.7 Motivations	11
1.8 Thesis Outline	14
Chapter 2 State of the Art of Research	17
2.1 Land Cover Classification using Remote Sensing	18
2.2 Image Classification Techniques	20
2.3 Tropical Forest Biomass Monitoring and Assessment	21
2.4 Definition of Biomass	22
2.5 Retrieval of Forest Biomass using RS Data	24
2.6 Biomass Model Transferability and Issues on the Model Accuracy	26
Chapter 3 Study of Geostatistics Texture Features for Complex Land Cover Mapping	29

3.1	Introduction	30
3.2	Satellite Image Preprocessing.....	31
3.2.1	Atmospheric corrections approaches	31
3.2.2	Radiometric calibration and topographic corrections methods	32
3.2.3	Spectral profiles and histogram of the ETM image.....	36
3.2.4	Image corrections results	37
3.3	Geostatistics texture features	38
3.4	Classification methods	41
3.4.1	Support Vector Machine Classification	41
3.4.2	Multi Layer Perceptron Neural Network	42
3.4.3	Variations of Input Data	42
3.5	Results and Discussion	43
3.5.1	Texture features generation.....	43
3.5.2	Image Classification	44
3.6	Discussion	47
3.6.1	Estimation of texture features moving window size.....	47
3.6.2	Behavior of texture features.....	48
3.7	Conclusions.....	52
3.8	Acknowledgments	52
Chapter 4 Improved Strategy for the Estimation of Above Ground Biomass using Non-Destructive Approach		55
4.1	Introduction	56
4.2	Data and Methods	58
4.2.1	Field data analysis	58
4.2.2	Satellite image preprocessing and vegetation indices generation	59
4.2.3	Estimation of stand volume	60
4.2.4	Above ground biomass (AGB) prediction	61
4.2.5	Spatial distribution of the AGB.....	62
4.3	Results and Discussion	63

4.3.1	Structure of forest stand properties	63
4.3.2	Prediction of stand volume.....	63
4.3.3	Conversion to above ground biomass.....	65
4.3.4	Spatial distribution of AGB	68
4.3.5	Relationship between AGB and land cover classification	70
4.4	Policy Implications	72
4.5	Conclusion.....	72
Chapter 5 Estimation of Forest Properties from Remote Sensing and GIS		75
5.1	Introduction	76
5.2	Data and methods	77
5.2.1	Field Observation Data	77
5.2.2	Images Acquisition and Preprocessing.....	77
5.2.3	Methods	79
5.2.4	Vegetation Indices Generation and Land Cover Classification	81
5.3	Results	82
5.3.1	Biomass Mapping using Field Data and GIS.....	82
5.3.2	Prediction and Dynamics of Biomass and Stem Volume using Remote Sensing	88
5.3.3	Comparison of AGB and stem volume estimates	93
5.3.4	Dynamics of biomass abundance	94
5.4	Discussion	96
5.4.1	Prediction Results Assessment	96
5.4.2	Relationship between GLCM Mean Texture, Land Cover, and Forest Biomass.....	98
5.5	Conclusions.....	99
Chapter 6 Retrieval of Forest Attributes using Bi-Temporal Remote Sensing Data.....		101
6.1	Introduction	102
6.2	Data and Methods	104
6.2.1	Image acquisition and preprocessing.....	104

6.2.2	Generation of ancillary remote sensing data.....	108
6.2.3	Field observation data	111
6.2.4	Statistical modeling.....	112
6.2.5	Validation	113
6.3	Results	114
6.3.1	Distribution of forest attributes.....	114
6.3.2	Correlations.....	115
6.3.3	Statistical models	116
6.3.4	Results evaluation.....	121
6.3.5	Forest attributes prediction from Bitemporal ETM data	123
6.3.6	Discussion	127
6.4	Conclusions.....	129
Chapter 7 Retrieval of Forest Biomass Using Mosaic SAR data.....		131
7.1	Introduction	132
7.2	Data and Methods	132
7.3	Results and Discussion	134
7.3.1	Estimation of AGB and Stem Volume based on field observation data	134
7.3.2	Relationship between radar backscatter and forest properties.....	135
7.3.3	Radar backscatter modeling.....	137
7.3.4	Prediction of AGB and Stem Volume using Radar Backscatter	138
7.3.5	Evaluation of AGB Estimate.....	139
7.4	Conclusions.....	141
Chapter 8 Dual-Polarimetry SAR and Optical Remote Sensing Data for Tropical Peatlands Discrimination		143
8.1	Introduction	144
8.2	Data and Methods	145
8.2.1	Preparation of satellite data and ground truth data collection	145
8.3	Methods.....	146
8.3.1	Matrix decomposition of SAR data	146

8.3.2	Image classification.....	149
8.4	Results and discussion.....	150
8.4.1	SAR Backscatter Responses.....	150
8.4.2	Canonical Discriminant Analysis.....	151
8.4.3	Classification of Radar Backscatter and Optical Data	155
8.5	Conclusion.....	156
8.6	Acknowledgement.....	158
Chapter 9 Summary		159
9.1	Thesis Contributions	159
9.2	Research Questions Answered	159
9.3	Conclusion and Recommendations	162
References		163
Appendix 1 Curriculum Vitae		179
Appendix 2 Publication List		182

List of Figures

Fig. 1.1 Deforestation rate in five major islands of Indonesia from 1985 – 2005 (MoF, 2006).....	3
Fig. 1.2 Logging roads along logged over secondary tropical forest in the Labanan concession (a) and left over timber from illegal forest harvesting caused wide gaps on forest canopies (b) (Photos by Yohanes Budi Sulistioadi).....	8
Fig. 1.3 Two study areas in tropical forests of Kalimantan, Central Indonesia, showing Labanan concession forest, a typical lowland <i>dipterocarp</i> forest in Indonesia (study site 1) and Tanjung Puting national park which comprises of tropical peatlands and swamp forests (study site 2) (Color figures is a composite of Landsat ETM image band 5, 4 and 3 in RGB channels).....	9
Fig. 1.4 Annual precipitation rate in the Labanan concession forest (source: Indonesian meteorological and Geophysics organization, 2007).....	10
Fig. 1.5 Peatland forest in Central Kalimantan Province (a) and destruction of peatlands due to forest clear cutting (b) (source: www.mongabay.com).....	13
Fig. 3.1 Spectral responses of multispectral bands Landsat ETM data for different land cover types (band 6 in the plot represents far infra red band of ETM band 7). The relative reflectance values clearly show the effects of atmospheric attenuations in partially corrected image (top of atmosphere reflectance profiles) as compared to surface reflectance profiles	35
Fig. 3.2 Histogram comparison of the uncorrected ETM image and corrected images based on a physical based MODTRAN method, and an image based DOS corrections approach. Both methods greatly reduced the attenuation effects in visual and near infra red bands and the histograms of the ETM image, either corrected using MODTRAN or DOS methods, show similar data distribution.....	37
Fig. 3.3 The result of topographic corrections in two subsets of the study site. The topography effects that cause shadows in the back side of the slope were reduced, although the study area is located in relatively flat region.	38
Fig. 3.4 Variogram plots of training data show the spatial variability of land cover classes under study.....	44
Fig. 3.5 Texture features derived from Landsat ETM data using spatial variability measures of Geostatistics and fractal dimension approach.....	45
Fig. 3.6 Accuracy assessments of classification results using different texture data. Each bar represents different variations of texture features used in the classification as explained in the figure legend.....	46
Fig. 3.7 Feature Space of rodogram and fractal dimension showing spatial distribution of total pixels on the study area and training data of each land cover class.....	50

Fig. 3.8 Subsets of maximum likelihood classification using the input data of (a) ETM 12347, (b) ETM 12347 and fractal dimension, (c) ETM 12347, fractal, variogram, pseudo-cross variogram, (d) ETM 12347, fractal, madogram, rodogram, pseudo-cross variogram, and (e) ETM 12347, fractal dimension, rodogram.....	51
Fig. 4.1 Plot of estimated and measured stand volume	65
Fig. 4.2 Comparison of biomass estimates. This study used stand volume – AGB equations, which include correction factors for dead AGB, below ground biomass. The study by Samalca and Brown estimated AGB from the allometric equation to relate dbh and AGB.....	66
Fig. 4.3 Comparison of biomass estimate with (a) basal area, (b) number of stems per ha, and (c) stand volume	67
Fig. 4.4 Kriging interpolation results (a) and error estimate (b).....	69
Fig. 4.5 Land cover map of the study area based on the interpretation of Landsat 7 ETM+ images and field survey (Modified from Berau Forest Management Project, 2001)	70
Fig. 4.6 Biomass estimate of each land cover class.....	71
Fig. 5.1 Workflow of the study describes two main approaches for estimating the AGB, using remote sensing method (left shaded box) and combination of field data and GIS method (right shaded box). The middle part of the workflow (non shaded area) explains the classification procedure of multi-temporal ETM images (2000 and 2003) performed in this study	80
Fig. 5.2 Plot of Stem volume/biomass and number of stems/ha vs. tree diameter (DBH)	85
Fig. 5.3 Landsat ETM of the study area with ETM bands 453 as RGB combination (a), and modified GIS land cover map of 2001 (b)	86
Fig. 5.4 Above ground biomass (AGB) and stem volume of each land cover type, sorted with the most advanced vegetation structures, i.e. mature forest-hilly, to the least complex structures, i.e. shrubs	87
Fig. 5.5 Classified Landsat ETM images of year 2000 (a) and 2003 (b) showing mature forest, very dense forests, dense forests, riparian forest, shrubs and bare soil. The bare soil class was masked out from the classification prior to the estimates of AGB density and stem volume of each land cover type	91
Fig. 5.6 Distribution of GLCM Mean Texture of Different Land Cover Type.....	98
Fig. 6.1 Distribution of forest properties in different diameter at breast height (DBH).....	115
Fig. 6.2 The linear regression and neural network estimations of basal area (a), stem volume (b) and biomass (c) against the observed data. The forest attributes include complete dataset (n = 338).....	121
Fig. 6.3 Comparison of multiple linear regression and neural networks predictions of basal area (a), stem volume (b) and biomass (c) against observed data by forest classes and z-test statistics of Wilcoxon signed rank sum test for paired	

dependent sample data. The regression and neural networks were tested separately against the observed test dataset (n = 112).....	123
Fig. 6.4 Comparison of bitemporal ETM data showing the subsets of ETM00 land cover classification (a) with the ETM00 above ground biomass (AGB) estimates (b), and the ETM03 land cover map (c) and the AGB predictions of the radiometrically calibrated ETM03 image (d).....	126
Fig. 6.5 Distribution of GLCM Mean Texture of Different Land Cover Type.....	128
Fig. 7.1 Estimates of above ground biomass and stem volume	135
Fig. 7.2 Correlation between (a) HH vs AGB, (b) HV vs AGB, (c) HV/HH vs AGB, (d) HH vs stem volume, HV vs stem volume, and (e) HV/HH vs stem volume.....	136
Fig. 7.3 Fitting of AGB model.....	138
Fig. 7.4 Fitting of stem volume model	139
Fig. 7.5 Landsat ETM data with band 4,5,3 in RGB channel (a) and the AGB estimate modeled from HV band (b)	140
Fig. 8.1 Dual polarimetry SAR data: (a) RGB combination of HH, VV, HH-VV bands, (b) scattering entropy, (c) anisotropy, and (d) alpha angle.....	147
Fig. 8.2 Radar backscatter, entropy, anisotropy, and alpha angle responses to different peatland classes.....	151
Fig. 8.3 Sample data plotted on the alpha angle / scattering entropy segmentation plane. In general, regions 1, 3 and 6 identify regions are dominated by multiple scattering, regions 2, 4 and 7 are dominated by volume scattering, and regions 5 and 8 characterize surface scattering mechanism.....	152
Fig. 8.4 Canonical discriminant function plot. The function 1 and 2 refer to the functions generated using HH, VV, entropy, anisotropy, and alpha as predictors, employing the coefficient of discriminant analysis	154
Fig. 8.5 Comparison of classification accuracy using confusion matrices for each class label, namely shallow peat – secondary forest (PSS), very shallow peat – sparse forest (PVSp), deep peat – primary forest (PDP), shallow peat – primary forest (PSP), palm oil plantation (PO), cultivated lands (CL), non-vegetated lands (NV) and water body (Wt)	155
Fig. 8.6 Maximum likelihood classification: (a) ETM Band 123457, (b) HH, VV, Alpha, Entropy, (c) ETM Band 123457, HH, VV, Entropy and (d) ETM Band 123457, HH, VV, Entropy-filtered with majority analysis. The water body (Wt) class label is masked out for better visual representation	157

List of Tables

Tab. 1.1 Materials of the study	6
Tab. 2.1 Recent studies on forest biomass assessment	23
Tab. 3.1 Variance matrix of forest cover classes training data based on Landsat ETM image.....	43
Tab. 3.2 Overall accuracy assessment (OAA) and Kappa statistics of spectral and geostatistics texture data classification.....	49
Tab. 4.1 Descriptive statistics of transect data.....	59
Tab. 4.2 Distribution of forest properties.....	63
Tab. 4.3 Prediction results of stand volume	64
Tab. 5.1 Descriptions of sampling plots describing parameters of different forest physiognomies.....	78
Tab. 5.2 Various simple ratios, different NDVIs, complex vegetation indices and image transform layers computed in this study to generate the biomass and stem volume equations	83
Tab. 5.3 Correlations between Remote Sensing Data, Stem Volume and Above Ground Biomass (AGB).....	89
Tab. 5.4 Comparison of stem volume and AGB from complete and selected datasets	90
Tab. 5.5 Classification Accuracy of Landsat ETM 2000 and 2003	92
Tab. 5.6 Comparison of AGB and stem volume estimates for particular land cover type.....	93
Tab. 5.7 Percentage of Land cover change from 2000 to 2003 based on Landsat ETM data classification (percentage is shown in brackets)	94
Tab. 5.8 Dynamics of Forest Biomass (AGB) and Stem Volume (Vol) from 2000 to 2003	95
Tab. 5.9 Above ground biomass estimates computed in this study using different allometric equations developed for tropical forest environment	97
Tab. 6.1 Description of sample data showing number of plots, total stems, mean diameter at breast height (DBH), basal area and stem volume for different forest physiognomies.....	106
Tab. 6.2 Various simple ratio, different NDVIs, complex vegetation indices and image transform layers computed in this study to generate the biomass and stem volume equations	109
Tab. 6.3 Correlations between remote sensing data and forest biophysical parameters	117

Tab. 6.4 Statistics and validation of linear regression analysis.....	118
Tab. 6.5 Statistics and validation of neural networks regression analysis.....	119
Tab. 6.6 Ordinary least square regression on randomly selected training pixels (n = 105844) and the root mean square error (RMSE) calculated from the testing pixels (n = 52921); α is the fitted intercept, β is the fitted slope, σ_α and σ_β are the standard deviation, and r is the correlation coefficient.....	124
Tab. 6.7 Comparison of mean intensities of randomly selected testing pixels (n = 52921) for the ETM 2003 before and after normalization against the image reference (ETM 2000), with paired t -tests for equal means (p -value = 0.01). The statistical tests were calculated from the uncalibrated ETM03 and ETM00 data	124
Tab. 6.8 Comparison of variances of the testing pixels (n = 52921) for the ETM 2003 before and after normalization against the image reference (ETM 2000), with F -tests for equal variances (p -value = 0.01). The statistical tests were calculated from the uncalibrated ETM03 and ETM00 data	124
Tab. 6.9 Forest attributes estimated from the sampling plots locations using the predictors of ETM00 and calibrated ETM03. The forest attributes were predicted using the regression models generated using the ETM00 data.....	125
Tab. 7.1 Statistics of forest biophysical properties.....	134
Tab. 8.1 Characteristics and carbon contents below the ground of different peatlands found in the study area.....	148
Tab. 8.2 Tests of Equality of Group Means	152
Tab. 8.3 Classification results and accuracy assessment based on cross validation ^(a,b)	154
Tab. 8.4 Accuracy assessment of the classification using ETM band 123457, HH, VV and entropy, and filtered with majority analysis using 3×3 moving window*	158

List of Abbreviations

AGB	Above Ground Biomass
ALOS Palsar	Advanced Land Observing Satellite/Phased Array type L-band Synthetic Aperture Radar
ARVI	Atmospherically Resistant Vegetation Index
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ATCOR	Atmospheric Corrections
AVHRR	Advanced Very High Resolution Radiometer
BA	Basal Area
BFMP	Berau Forest Management Project
BGB	Below Ground Biomass
BMG	Badan Metereologi dan Geofisika (Indonesian Metereological and Geophysical Institution)
BRDF	Bidirectional Reflection Distribution Function
CIRAD	Center of International Cooperation in Agronomic Research for the Development
dbh	Diameter at Breast Height
DEM	Digital Elevation Model
DN	Digital Number
DOS	Dark Object Subtraction
EOS	Earth Observation Satellite
ETM	Enhanced Thematic Mapper
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FWI	Forest Watch Indonesia
GEMI	Global Environmental Monitoring Index
GFW	Global Forest Watch
GIS	Geographic Information System
GLCM	Grey Level Co-occurrence Matrix
Gt	Gigaton
IDL	Interactive Data Language
JAXA	Japan Aerospace Exploration Agency
JERS-1	Japanese Earth Resources Satellite, number 1
Lidar	Laser Imaging Detection and Ranging
LULUCF	Land Use, Land Use Change and Forestry
MAD	Multivariate Alteration Detection
Mg	Megagram
MIR	Middle Infra Red
MLC	Maximum Likelihood Classification
MLP-NN	Multi Layer Perceptron Neural Network
MODIS	Moderate Resolution Imaging Spectroradiometer
MoF	Ministry of Forestry

MONCER	Design and Development of an Effective Monitoring and Certification System to Support Sustainable Management of Production Forest in Indonesia
MSAVI2	Modified Soil Adjusted Vegetation Index 2
MSE	Mean Square Error
MSS	Multispectral Scanner
NDVI	Normalized Difference Vegetation Index
NIR	Near Infra Red
NSEW	North-South-East-West
OAA	One Against One
OAA	One Against All
OAA	Overall Assessment Accuracy
PCA	Principal Component Analysis
Polinsar	SAR Polarimetry and Polarimetric Interferometry
Radambrasil	Radar Mapping of Brazil
RGB	Red Green Blue
RMSE	Root Mean Square Error
RS	Remote Sensing
RTRWP	Rencana Tata Ruang Wilayah Propinsi (Provincial Land Use Planning)
RVoG	Random Volume over Ground
SAR	Synthetic Aperture Radar
SAVI	Soil Adjusted Vegetation Index
SD	Standard Deviation
SEE	Standard Error Estimate
SIR-C	Spaceborne Imaging Radar, C-Band
SR	Simple Ratio
SRTM	Shuttle Radar Topography Mission
SSC	Single Look Slant Range Complex
STREK	Silvicultural Techniques for the Regeneration of Logged Over Forest in East Kalimantan
SVM	Support Vector Machine
TC	Tasseled Caps
TM	Thematic Mapper
TOA	Top of Atmosphere
TSP	Temporary Sampling Plot
UNFCCC	United Nations Framework Convention on Climate Change
UTM	Universal Transverse Mercator
VI	Vegetation Index
WGS	World Geodetic System

Chapter 1

Introduction

ABSTRACT

The research presented in this thesis deals with the development of alternative approaches to improve the classification accuracy of remotely sensed data and to propose an alternative non-destructive approach for mapping of above ground biomass over tropical forest landscape. This chapter introduces the background of the study, explaining tropical forest disturbance and monitoring and deforestation in Indonesia and its impacts on carbon storage. Also, research questions, scientific challenges and objectives, materials and methods used for the study, description of the study areas, motivations behind the present study, and a thesis outline are briefly explained.

Keywords: introduction, background, research questions, scientific challenge, objectives, materials and methods

1.1 Tropical forest disturbance and monitoring

Forest is a specific territory of ecosystems and consists of vegetation formation with a continuous canopy of trees (FAO, 2003). Forest ecosystems are distinguished depending on the dominant vegetation type, stand structure, climate, soil type, and topography. Based on the local climate forest ecosystems may be divided as tropical, boreal and temperate forests. In the last two decades, tropical forest is the most severely degraded and deforested compared to other forest ecosystems, mainly due to anthropogenic factors and natural hazards. Tropical rainforests belong to the most heterogeneous ecosystems and exhibit the highest biodiversity in the world, comprising more than half of all the world plant and animal species (Butler, 2006). Although the tropical rainforests cover only less than 6% of land surface on the earth, they play an important role for the global carbon cycle, as tropical rainforests store 30% of all terrestrial carbon (up to 200 Mg/ha) and sequester about 12 Mg C/ha per year, i.e. 22% of all carbon fixed on Earth. Carbon dioxide is released from forest ecosystems during decomposition process or when the forest is disturbed or destroyed, either due to natural hazards or human activities. Tropical forest ecosystems recently have been severely devastated, mainly due to excessive forest encroachments, such as timber harvesting, conversion to agricultural land and settlement area, and exploitation of non-timber forest resource (Collins et al., 1991).

Monitoring of forest resources is essential to assess their carrying capacity and to observe the dynamics of forest cover changes. Mapping of forest cover is an ultimate way to assess forest cover changes and to study the forest resource within a period of time (Wijaya et al., 2008b). Estimation of forest biophysical parameters, such as stand volume and woody biomass is also important for forest inventory and management and for scientific purposes (Parresol, 1999).

1.2 Deforestation in Indonesia and its Impacts on Carbon Storage

Indonesia possesses one of the largest tropical forests after the Amazonian in Brazil and the tropical rainforest of the Republic of Congo, representing 10% of the remaining tropical forests in the world (FWI/GFW, 2002). With total terrestrial areas of 187.9 million hectares, the Indonesian Ministry of Forestry (MoF) noted that 133.7 million hectares or more than 70% areas were designated as forests (MoF, 2006). These do not include inland water ecosystems which cover 3.4 million hectares. Tropical forests in Indonesia support very high biodiversity and are considered among the richest ecosystems in the world. More than 16% of the entire world's bird species, 11% of plant species, and 10% of all mammals are

found in the Indonesian tropical forests. Among the unique species found in this region are the endangered species orangutan and Sumatran tiger.

Compared to other South East Asian countries, Indonesia has the highest deforestation rate, and this is of major local and global concern (MoF, 2002). We noted that about 1.87 million ha/year of Indonesian forests were deforested from 1985 to 1997. This situation worsened during a severe drought event of El Nino in 1997 – 1998 causing sporadic and devastating fires in most tropical rainforests in South East Asian region, especially in Indonesia (Page et al., 2002). After the El Nino drought event in 1998, the government updated the database recording forest losses in five main islands of Indonesia, and reported that deforestation rate from 2000 – 2005 was over 1 million ha/year, whereas in 2002 – 2003 the deforestation rate reached more than 1.9 million ha (Fig. 1.1).

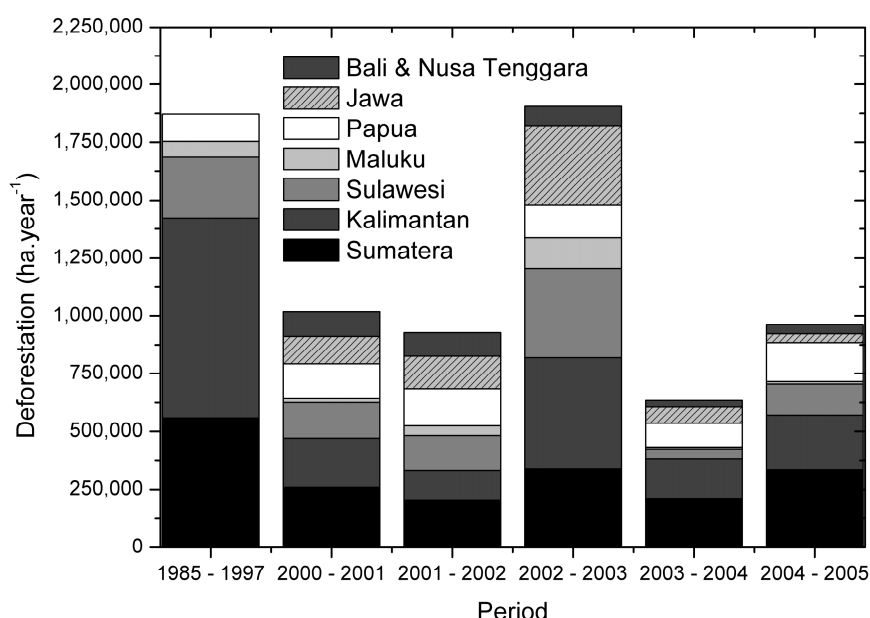


Fig. 1.1 Deforestation rate in five major islands of Indonesia from 1985 – 2005 (MoF, 2006)

Deforestation, peat degradation and forest fires are major contributors of carbon emissions in Indonesia (Sari et al., 2007). The amount of carbon released from deforestation and forest fires is five times greater than that of the non forestry sector, and the number of carbon released from the forestry sector has made Indonesia as one of the top five carbon emitters in the world (Sari et al., 2007). Of about 19 billion tons of terrestrial carbon stocks (80% of all terrestrial carbon) in Indonesia are stored in the forested lands, and a single devastated forest fire, such as El Nino event in 1997, had released 0.19 – 0.23 Gt (1 Gigaton = 1 billion tons) of carbon to the atmosphere only from peatland forests (Page et al., 2002). If this number was extrapolated for the entire country, 0.81-2.57 Gt of carbons or equal to 13-40%

of mean annual global carbon emission from fossil fuels were liberated in 1997 (Hoekman, 2007).

We consider it as scientifically challenging and desirable to develop accurate and adequate means of forest cover mapping and modeling the forest biophysical parameters, such as biomass, over larger areas. The research proposed here aims at the development of remote sensing and geographical information systems, which are relevant for the mapping of forest cover and the estimation of forest biophysical parameters. The application of remote sensing data is the only possible way for effectively and efficiently extracting this information over large areas.

1.3 Research Questions

Several research questions are to be answered in this present work and explained as follows.

1. Which texture features are useful to improve the classification accuracy over tropical forest landscape?
2. How can multi-source remote sensing coupled with in situ data improve the assessment of above ground biomass (AGB)?
3. Which types of remotely sensed data and vegetation indices are important for the modeling of the AGB?
4. How do land cover changes affect the dynamics of forest biomass?
5. How can AGB changes be estimated using bi-temporal satellite data? What is the predictive ability of linear regression method as compared to a non-parametric neural network model?
6. Which features generated from Synthetic Aperture Radar (SAR) data are more important to assess the biomass? How accurate is the estimated biomass model?
7. How does the combination of Landsat ETM sensor and SAR data improve the classification of peatland forest?

1.4 Scientific Challenge and Objectives

The objectives of this thesis are:

1. To briefly review already published approaches for the mapping of forest cover and modeling of forest parameters, above ground biomass (AGB) in particular, defining the limitations of these approaches and identifying possibilities for innovation in mapping the forest biomass distribution.

2. To investigate the possibility of combining texture data with Landsat ETM data for improving the accuracy of land cover classification in complex forest landscape.
3. To develop a non-destructive approach to model the spatial distribution of above ground biomass applying the spectral information and vegetation indices, modeled using neural networks method and kriging interpolation technique.
4. To develop a model to estimate the above ground biomass (AGB) using Landsat ETM data, vegetation indices, and digital elevation models (DEM), and to study the possibility to assess the AGB from a bi-temporal dataset.
5. To investigate the potential of mosaic ALOS Palsar data for AGB modeling and to study the possibility for mapping the AGB over larger forest regions.
6. To combine the Landsat ETM data and SAR data for land cover classification in tropical lowland- and peatland forests, and to assess the drawbacks of the data combination.
7. To make an integrated assessment of the complexity of deforestation and forest degradation in relation to the forest cover dynamics and the modeling of the AGB over a tropical forest environment.

1.5 Materials and Method

This thesis is generally divided into two main topics: (1) land cover classification in complex forest landscape using optical data and geostatistics texture features and (2) estimation of biophysical forest parameters employing optical data and synthetic aperture radar (SAR) image. Besides, a land cover map, precipitation data and digital elevation models were also collected as additional information (Tab. 1.1).

The Landsat ETM data preprocessing was conducted beforehand. Corrections of atmospheric and topographic effects were carried out using a standard method explained in Chapter 3. The SRTM DEM was resampled into a 30 m resolution to fit with the ETM data. The topographical map, land use map, and SAR data were georeferenced with the ETM data and projected into UTM coordinate system and WGS84 datum. Descriptions of the satellite data corrections, image preparation, filtering, classification technique, and the modeling of biophysical parameters from RS data are briefly explained further in each chapter.

Tab. 1.1 Materials of the study

Data type	Acquisition date	Path/Row	Spatial resolution	Season	Description
Landsat ETM	31 May 2003	117/59	30 m	Beginning of dry season	Clear atmosphere, with almost no clouds
Landsat ETM	26 August 2000	117/59	30 m	Dry season	Hazy conditions, some clouds persist
Mosaic ALOS Palsar	2006 - 2008		50 m		Image preprocessing was conducted by JAXA
SRTM DEM		117/59	90 m		Downloaded from USGS website
TerraSAR X data	13 March 2008		6.5 m	Mid wet season	Single Look Sland Range Complex (SSC) image, dual-polarization (HH and VV)
Land use/land cover map of Kalimantan	2002				Generated from 1:50,000 topography map and observation on the ETM data (National Surveying Agency of Indonesia)
Peatlands map of Kalimantan	2004				Provided by Wetlands International – Indonesia Programme
Precipitation data					Obtained from Indonesian meteorological and Geophysics organization (Badan Metereologi dan Geofisika/BMG)
Forest inventory data	April 1997 – March 1998				Circular sampling plot, data was collected in transect.
Field work data	September 2004				Circular sampling plot, data was collected based on purposive sampling
Ground truth data	March 2007				Collecting additional ground truth data

1.6 The Study Areas

1.6.1 Labanan concession forest (study site 1)

Chapter 3, 4, 5 and 6 of this thesis consider the Labanan concession forest in Berau District, East Kalimantan Province, Indonesia. The concession area consists of mostly logged over/secondary forests mixed with some primary forest patches (Fig. 1.2). Several research projects have been conducted in this forest concession.

Silvicultural Techniques for the Regeneration of logged over forest in East Kalimantan (STREK) Project, a cooperation between the Center of International Cooperation in Agronomic Research for Development (CIRAD) and the Indonesian Ministry of Forestry (MoF) was conducted from 1989 to 1996 to develop silvicultural and management rules for sustainable forest productivity (Fauzi, 2001). Berau Forest Management Project (BFMP) was established from 1996 to 2002 as a cooperation between the European Union and the MOF of Indonesia continuing the related research to support sustainable forest management (BFMP, 1997). The Ministry of Forestry gave a special status to Labanan concession in 1999 as a research forest, which was then certified by ISO 14001 in 2001. Currently, design and development of an effective monitoring and certification system to support sustainable management of production forest in Indonesia (known as the MONCER project) is operating. This project aims to design and to implement a practical, purpose-oriented, effective and efficient forest certification system for supporting a monitoring system of sustainable forest management (Widayat, 2005).

The study area is geographically situated along the equator at the coordinates of 1°45' to 2°10' N, and 116°55 to 117°20' E and has a size of 83,240 hectares, of which 54,567 ha are production forests, 26,997 ha are limited production forests, and 1,676 ha are non-production forests (Fig. 1.3). According to Provincial land use planning (Rencana tata ruang wilayah propinsi: RTRWP), 81,564 ha were allocated for forest areas (kawasan budidaya hutan) and 1,676 ha were allocated for non-forest areas (kawasan budidaya non-kehutanan) (Wijaya et al., 2005).

The forest area is mainly situated on inland of coastal swamps and consists of undulating to rolling landforms with isolated masses of highly and high mountain (BFMP, 1997). The lowland landscape is derived from sediments, mudstone, siltstone, sandstone and gravel, laid down in the Miocene and Pliocene periods, and recent alluvial deposits.



(a)



(b)

Fig. 1.2 Logging roads along logged over secondary tropical forest in the Labanan concession (a) and left over timber from illegal forest harvesting caused wide gaps on forest canopies (b) (Photos by Yohanes Budi Sulistioadi)

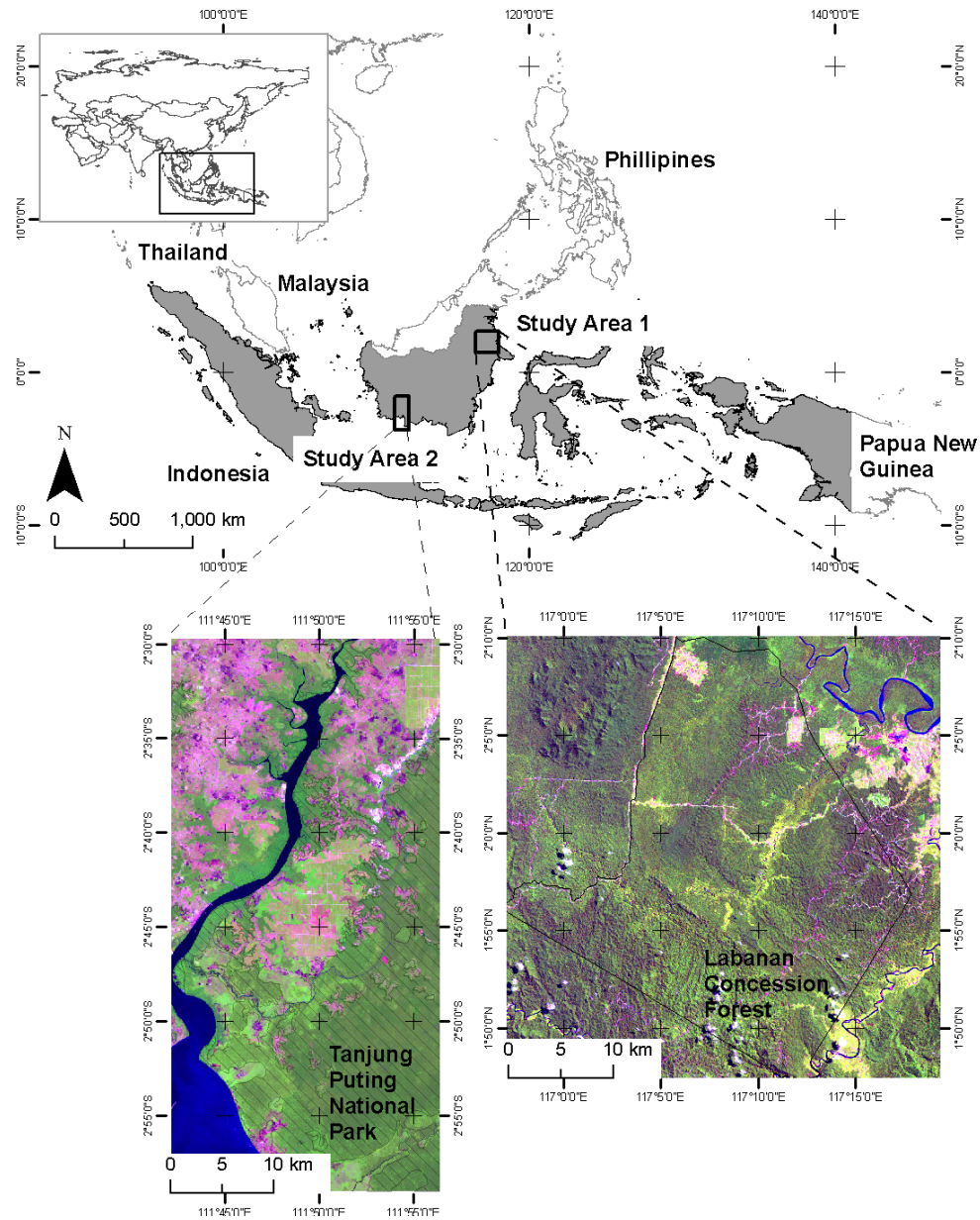


Fig. 1.3 Two study areas in tropical forests of Kalimantan, Central Indonesia, showing Labanan concession forest, a typical lowland *dipterocarp* forest in Indonesia (study site 1) and Tanjung Puting national park which comprises of tropical peatlands and swamp forests (study site 2) (Color figures is a composite of Landsat ETM image band 5, 4 and 3 in RGB channels)

Significant erosion has developed a landscape which is relatively low lying, but broken. Relative relief is seldom more than 50 m, although the elevation reaches almost 500 m above sea level in the hills to the west of the Labanan concession. Wetlands are not particularly important in the Labanan area with only limited swampy areas along the Segah and Kelai rivers. Soils are generally well drained, red clayey and highly leached, i.e. poor in nutrients, tropudults or paleudults (red-yellow podzolics in the Indonesian classification), but with a relatively high silt and fine sand content. The soils are highly erosion prone. Land slides are common on road cuttings and on steeper uncleared land. Slopes in the region are generally rolling to steep and very steep in places, though slope lengths are not often long. Selective felling of the natural forest with reforestation of degraded areas is the best use for most of the forested land (Mantel, 1998).

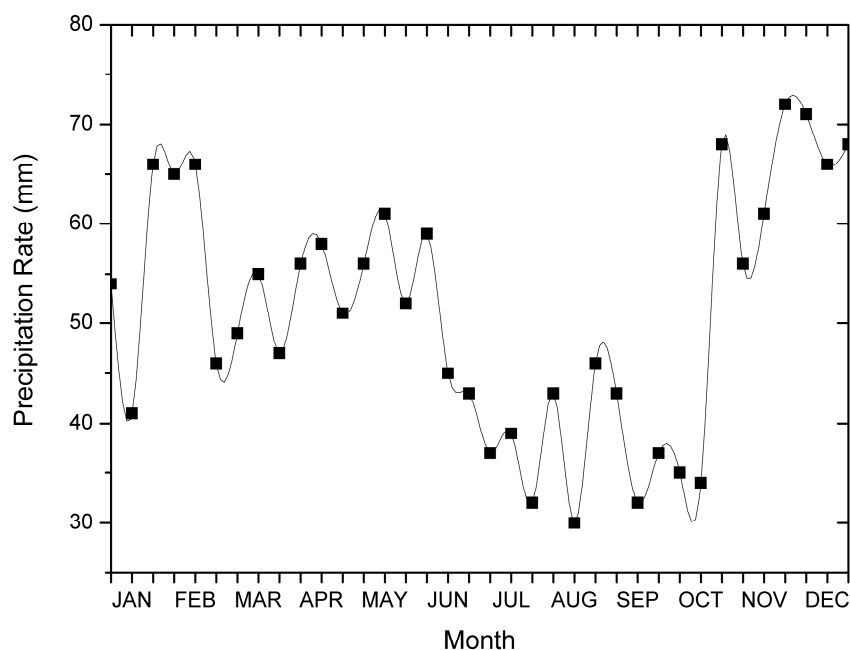


Fig. 1.4 Annual precipitation rate in the Labanan concession forest (source: Indonesian meteorological and Geophysics organization, 2007)

The climate is wet tropics with rainfall ranging from 1800-2000 mm per year on average at Tanjung Redeb (capital of Berau municipality), but rising rapidly inland to almost 3000 mm (BFMP, 1997). The months of June to October are drier; the wettest months are from November through to February. No months receive less than 30 mm on average, though extended dry spells do occur they are not annual or regular (Sist and Nguyen-Thé, 2002). The average annual and monthly precipitation rates in the Labanan concession are 1840 mm and 51 mm, respectively (Fig. 1.4). Temperatures are between 23 - 33 degrees Celsius with a long term average of 26 degrees Celsius. Relative humidity is 91% or higher and

winds seldom exceed 9 km/hour. The forest is a mixed lowland *dipterocarp* forest with Shorea, Dipterocarpus and Vatica as the predominant genera in the pristine forest followed by the Euphorbiaceae family (BFMP, 1997).

1.6.2 Tanjung Puting national park (study site 2)

Chapter 7 of this thesis considers another study area over peatland forests in Central Kalimantan Province. The Borneo is actually the third largest island in the world that is comprised of Malaysian and Brunei Darussalam territories and for the most part the Indonesian provinces of Kalimantan. The Kalimantan is approximately 543,900 km² or 73% of Borneo and has a population of 11.3 million (Hecker, 2005). The Indonesian Borneo is divided into four provinces: Kalimantan Barat (West Kalimantan), Kalimantan Tengah (Central Kalimantan), Kalimantan Selatan (South Kalimantan), and Kalimantan Timur (East Kalimantan).

Peatland forests in Central Kalimantan are one of the few regions in the world, where unique and valuable tropical peatlands can be found. A peat layer of up to 20 meters deep mostly covers this area of more than 1 million hectares of peatlands in this province (Wetlands International, 2007). Unfortunately, a large portion of the peatlands is severely degraded and only small peat swamp forests remain. Nowadays, the degraded peatlands which were used for agriculture, industrial plantations, and settlements, are mostly abandoned. As a consequence, these areas continuously suffered from major forest fires and river floods (Wetlands International, 2007).

This study considers the Southern part of Central Kalimantan province, which is one of the most problematic areas in terms of peatland degradation in SE Asia (van Beukering et al., 2008). The study site includes Tanjung Puting National Park, which has a total size of 400 thousand hectares and is mostly dominated by tropical heath forest, peat swamp forest and mangrove forest (Fig. 1.5). This national park is recognized as one of the world biosphere reserves by the United Nations and forms the largest protected peat swamp forests in the SE Asia. It is situated between 2°35' – 3°20' South and 111°50' – 112°15' East (Fig. 1.3) (Susilo, 1997). The national park has been highly degraded and deforested at an alarming rate, mainly due to illegal forest harvesting, opening of agricultural lands, and conversion to palm oil plantations.

1.7 Motivations

A considerable amount of research has been done in improving the accuracy of land cover classification and the modeling of biophysical properties over forested regions using remote

sensing data and approaches. Nevertheless, the following encourages this study which mainly has an ultimate goal to enrich scientific knowledge and references of this particular subject.

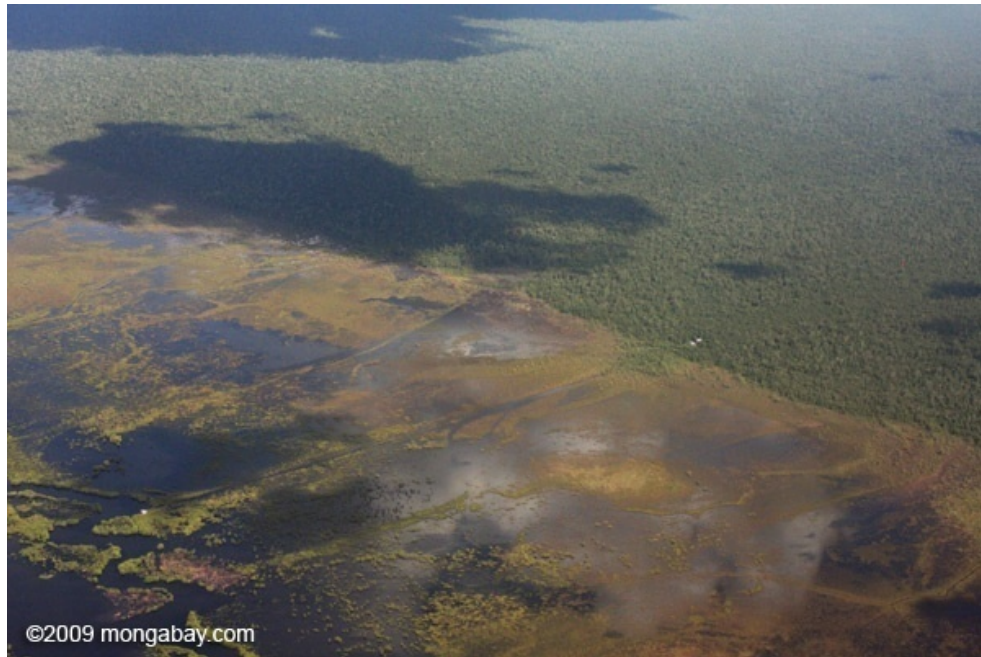
Transferability problem of available biomass models

Transferability of biomass model developed in tropical forest environments and estimated from remote sensing data is rather limited (Foody et al., 2003). This is because the allometric biomass models generated from a particular forest landscape cannot be easily applied for another forest region. We, therefore, are encouraged to generate the biomass model which is suitable for tropical forest environment in Indonesia, as well as to compare its performance with the existing model estimated from other forest regions.

To develop an alternative approach in combining optical and microwave sensors for land cover classification

Data fusion is the seamless integration of data from distinctive sources. Optical and microwave RS are complementary to each other as their characteristics are different. Optical data basically measures the physical properties of ground objects, whereas the SAR data provides more information about geometric properties of the objects (Lillesand and Kiefer, 1994). Moreover, the capability of microwave data to penetrate the atmosphere under virtually all conditions can also give an advantage for the fused data.

The fusion of both data can be based on pixel-level or feature-level image fusion. Pixel-level image fusion combines pixel by pixel of the reflectance of optical images and radar backscatter to generate the fused image, which has more information about object features (Wu and Chen, 2007). Feature-level image fusion on the other hand, is conducted by combining both sensors in the classification processes employing the neural networks method (Kuplich, 2006), or by linking the optical and SAR data classification results applying Dempster-Shafer evidence theory (Wu and Chen, 2007). Another alternative is to fuse optical and microwave sensors using a Bayesian approach which jointly uses the classification results of individual sensors to improve class labeling using statistical measures (Venkataraman et al., 2004). The fusion of optical and SAR data has recently been of great interest for land cover classification because SAR data are widely available, and due to the fact that the data fusion is promising resulting in more accurate information of the classified image.



(a)



(b)

Fig. 1.5 Peatland forest in Central Kalimantan Province (a) and destruction of peatlands due to forest clear cutting (b) (source: www.mongabay.com)

To enrich scientific studies on biomass over another tropical rainforests environment beside the Amazon

A great number of studies, as exhibited in Tab. 1.1, has been conducted in the tropical forest of the Amazon, and only few studies were carried out in another tropical forest landscapes, such as in Indonesia. The forested lands in this region were heavily deforested at an alarming rate and greatly devastated especially after the severe El Nino drought and mass forest fires in the 1990's (Stibig and Malingreau, 2003). Few references are found conducting forest biomass studies in the South East Asia region especially those implementing remote sensing data (Foody et al., 2001). We found even fewer studies conducting field measurement approach for estimating forest biomass in Indonesia (Samalca, 2007).

This motivates us to undertake this study in order to enrich the knowledge and scientific references on biomass estimation and modeling over tropical forest regions, especially in Indonesia.

1.8 Thesis Outline

In chapter 2, the state of the art of research is briefly presented. A literature review of some recent publications on land cover classification is presented first, followed by a review of different image classification techniques applied in remote sensing. The review on forest biomass monitoring and assessment is explained afterwards, followed by some recent studies on the retrieval of forest biomass using remote sensing approaches. Finally, the issues of biomass model transferability and model accuracy are reviewed.

Chapter 3 of this thesis focuses on the classification of Landsat ETM data and geostatistical texture features to increase the classification accuracy of complex land cover mapping. We compared the conventional maximum likelihood method, support vector machine (SVM) classification, and neural networks algorithm. Different variations of the texture features were tested to find the best input combination for the classification.

The study proposing a non-destructive method for estimating above ground biomass (AGB) from remote sensing is presented in chapter 4. Landsat ETM data combined with number of stems are used to model the AGB using neural networks algorithm and kriging method. The estimation is compared with recent studies and the results are discussed.

Chapter 5 presents the utility of remote sensing, GIS and field observation data to model AGB and stem volume, as well as their changes over time. Ancillary data from the Landsat ETM bands reflectance, e.g. vegetation indices, PCA bands and GLCM texture, are generated and used for the modeling. Two land cover maps are generated from the Landsat ETM 2000

and 2003 to predict the AGB over different forest classes. The biomass level and its changes over time are then estimated.

In chapter 6, the quantification of forest parameters in different successional forest stages is investigated. This information is useful for understanding the relationship between different regenerating stages and ecosystem changes. For the modeling of forest attributes, the multiple linear regression technique and neural networks approaches were used and compared. A possibility to estimate the forest attributes from bitemporal Landsat ETM data was conducted by normalizing the second Landsat ETM data using the multivariate alteration detection (MAD) method.

The retrieval of AGB and stem volume from SAR data is experimented in chapter 7, where mosaic ALOS Palsar data is employed to model both forest properties. The radar backscatter and their ratios are used to model the forest properties employing non-linear models. The results are discussed afterwards.

Integration of SAR backscatter and optical data for peatlands classification is presented in chapter 8. The TerraSAR X data and Landsat ETM data are used to discriminate peatlands in tropical swamp forest of Central Kalimantan. Discriminant analysis is used to select the input data which are useful for the classification. Using a maximum likelihood classification technique, the fusion of SAR and Landsat ETM data is conducted and the accuracy is assessed using confusion matrices.

Chapter 2

State of the Art of Research

ABSTRACT

This chapter presents a brief review on the application of remote sensing data approaches for land cover mapping and forest parameters modeling and estimation. It begins with a brief description on land cover classification using remote sensing data and approaches, followed by a description of image classification techniques. Subsequently, reviews on tropical forest biomass monitoring and assessment are explained. Some theoretical and practical definitions of biomass are also described, including the definition of biomass used for this particular study. Furthermore, references related to the retrieval of forest biomass using remote sensing data and the limitation of biomass model transferability and issues on the model accuracy are presented.

Keywords: image classification techniques, model transferability, forest, land cover classification, remote sensing, biomass

The topics of land use, land use change and forestry (LULUCF) have become more critical issue for Indonesian rainforests for over the last two decades, since a large number of forests in the country has been severely degraded, deforested, and/or converted into another land use, such as for cultivation, urban areas, mining areas, etc. Forest management practices for supporting sustainable forest management greatly depend on the development of new methods and techniques based on scientific approaches. The importance of remote sensing data for forest inventory and management has increased since the conclusion of the first meeting related to global change issues, which was mainly encouraged by the discussions on carbon emissions explaining a huge number of mature tropical forests conversion into different landscape and/or various secondary forest stages (Ramankutty et al., 2007). This chapter explains the state of the art of the research of remote sensing applications, mainly focusing on land cover classification and biomass modeling and estimation.

2.1 Land Cover Classification using Remote Sensing

Provision of accurate information on land use/land cover requires an accurate land cover map, mainly generated from the interpretation and classification of remote sensing data. Remote sensing and digital image processing research have produced many approaches to multi-spectral land cover classification over the last two decades, with major differences in accuracies attributed to both training methods and classification algorithms (Yool, 1998).

Optical sensors, e.g. Landsat data, are probably the most popular sensors for land cover classification because the data has relatively high temporal resolution and is widely available. Wide variations of spatial resolutions, e.g. 15m (ASTER), 30m (Landsat), 250m and 500m (MODIS), and 1km (AVHRR), have also given flexibility to the data applicability. Upon image corrections and preprocessing, the classification employing these data are relatively straightforward, and the complexity of the classification procedure greatly depends on the selected classification algorithm and methods. Spectral information (normalized digital number or DN) of the satellite data is usually used as input for the classification (Wijaya et al., 2008b).

Vegetation indices (VI's) are commonly used as an additional input of spectral data to increase the discrimination of different forest classes or forest disturbances. Phenological changes of different tree species were captured using the conjunction of Landsat MSS and TM data, normalized difference greenness index (NDGI), and normalized difference vegetation index (NDVI) (Wolter et al., 1995). The indices, which can be generated from single bands, band ratios, vegetation indices and multivariate components of the Landsat

data, are also useful to assess burn severity due to forest fires (Epting et al., 2005). Despite of its advantage, the sensitivity of vegetation indices is rather limited to some wavelength windows, such as red or near infra red bands, and the saturation problem which commonly occurs for some VI's also reduces their applications (Huete et al., 1997).

In addition to spectral information, the utility of spatial information in the form of texture features may also be useful for the image classification. The inclusion of texture information involves the information from neighboring pixels which is important to characterize the identified objects or regions of interest in an image (Haralick et al., 1973). The texture data is particularly important if the classification involves no dominant distinct object (i.e. object with particular shape or boundary) and spatial correlation of neighboring pixels of that particular object forms similar texture variation (Jakomulska and Clarke, 2001). This study experimented with the classification of texture layers generated using geostatistics approach, as demonstrated in several recent studies (Chica-Olmo and Abarca-Hernandez, 2000; Miranda and Carr, 1994; Miranda et al., 1998; Zhang et al., 2004). These studies found that the classification accuracy is notably improved when the geostatistics texture layers were used as additional input for spectral data classification. Additionally, the inclusion of texture features calculated from fractal dimension is also experimented in the present study.

Application of synthetic aperture radar (SAR) data for land cover classification is greatly encouraged by the problem of optical data particularly in humid and tropical regions due to the presence of persistent clouds (Salas et al., 2002). The utility of multi-wavelength SAR data was demonstrated by Kuplich et al. (2005) combining JERS-1, SIR-C and XSAR and optical bands from Landsat TM for discriminating regenerating forest stages in the tropical Amazon. Classification of flooded and upland vegetation in tropical forest landscape was also conducted with success using L-band JERS-1 data (Miranda et al., 1998). Recent generation of ALOS Palsar data has been planned to cover global tropical forest environments to frequently map the forest cover over these regions. This is initiated by the Kyoto and Carbon Initiative Project and fully supported by JAXA, the provider of ALOS Palsar data (Rosenqvist et al., 2008). However, the great challenge in remote sensing study recently is to combine optical and microwave (i.e. SAR data) sensors for improving the classification accuracy and to generate better visual characteristics of the classified image (Lu, 2006). Attempts to combine optical and SAR data are also experimented in the present study, combining both sensors for the classification of peatlands in tropical swamp forest and to discriminate forest regenerating stages in a tropical lowland forest regime.

2.2 Image Classification Techniques

Multispectral image classifications are basically attempts for finding patterns in the spectral response in relation to land cover groups known to be present. The thematic map describing forest characteristics is often the goal of an image classification in a forestry context (Wulder, 1998). Image classifications label pixels into classes, or categories, based on distinctive patterns of digital numbers (DN value). The classification procedures are normally grouped either as supervised or unsupervised. For more accurate information, a supervised classification is recommended as it involves field observation data and the classification has a strong basis on a priori knowledge of land cover over a study area (Lillesand and Kiefer, 1994).

The most popular method for image classification is the maximum likelihood approach. This approach is a statistical decision rule method that examines the probability of a pixel in relation to each class with assignment of the pixel to the class with the highest probability (Jensen, 1996). It has an underlying assumption of a normal distribution to the data within each class and may be biased with unequally sized training classes.

Another approach which needs no assumption on data distribution is the neural networks method. Neural networks are essentially learning systems based on interconnected networks of simple processing elements (Rumelhart et al., 1986), and when applied for image classification this method is concerned with the transformation of data from feature space to class space. The wide application of neural networks in remote sensing is due mainly to their ability to perform more accurately than other techniques such as statistical classifiers, particularly when the feature space is complex and the source data has different statistical distributions (Atkinson and Tatnall, 1997). Various neural network types have been widely studied and explored in remote sensing studies mainly for satellite image classification. Feed forward back-propagation neural network is probably among the most popular neural network approaches applied for the classification of multi-source remote sensing data (Kanellopoulos and Wilkinson, 1997). This particular neural network is proven effectively to enhance the classification of forest regenerating stages (Liu et al., 2005), although the training algorithm is relatively time consuming and sometimes fails to find the optimum solution (Ardo et al., 1997). More effective learning algorithms are explored to reduce the network training, like combining Kalman filter and conjugate gradient methods in training the network (Canty, 2009; Marpu et al., 2008). The development of an automatic mapping of land cover classification using fuzzy ARTMAP neural networks by finding the minimal number of recognition categories to meet the accuracy criteria was also studied (Carpenter et al., 1997). The possibility to increase the spatial resolution of a land cover map

at sub pixel level by combining the information from the land cover proportion data and fused imagery was explored by Tatem et al. (2001).

Another non-parametric classification approach is support vector machine (SVM) which is based on statistical learning theory (Vapnik, 2000). The learning theory was developed for solving pattern recognition problems. The SVM, originally introduced by Vapnik (1999) is a classification and regression technique which is now widely used in very different fields, including remote sensing (He et al., 2005). Recent remote sensing studies evaluated the performance of the SVM for land cover classification and compared the results with more popular classification techniques, like maximum likelihood and neural networks classification (Hahn et al., 2007; Huang et al., 2002; Pal and Mather, 2005). Those studies show that SVM has the potential to increase classification accuracy. This study, moreover, applies the SVM for image classification in the tropical rainforest landscape combining Landsat ETM data and texture information generated from geostatistics and fractal dimension algorithm (Wijaya and Gloaguen, 2007b; Wijaya et al., 2008b). In general, the SVM method is relatively faster than standard neural networks classification. However, the main challenge of this method is to select ground truth data which optimize class separability from the predefined hyperplane.

2.3 Tropical Forest Biomass Monitoring and Assessment

Assessment of forest biomass has two major objectives: (1) for resource use and (2) for environmental management (Parresol, 1999). The first objective implies that to determine how much fuel wood or timber available for use is important as we need to know how much biomass is available at any given time. In environmental management, biomass quantification is important to assess the productivity and sustainability of forests. Biomass and its temporal changes is also an important indicator for carbon sequestration estimation. For this purpose, we need to know how much biomass is lost or accumulated over time. Since over 50% of the forest dry biomass is carbon, the amount of carbon sequestered can be inferred from the biomass change (Losi et al., 2003). The Kyoto protocol requires transparent reporting of forest removal and accumulation (biomass change). This implies we need the development of precise procedures to quantify forest biomass, which generally can be divided into above ground- and below ground biomass. Above ground biomass (AGB) consists of all living and dead biomass above the soil including stem, stump, branches, bark, foliage and seeds or fruits. The below ground biomass consists of all living and dead roots. Fine roots of less than (suggested) 2 mm diameter are excluded because these often cannot be distinguished empirically from soil organic matter or litter (FAO, 2004).

There are three approaches to mapping biomass in forest areas, which are field measurement, remote sensing, and GIS-based approaches (Lu, 2006). Field measurement is considered to be accurate but proves to be very costly and time consuming as destructive sampling is required (De Gier, 2003). While the GIS approach is not widely studied so far, the remote sensing approach has become an efficient and popular technique for many ecological studies, including for biomass assessment taking into account the launching of several new sensors into space with more sophisticated technology (e.g. increased temporal, spectral and spatial resolutions). Recent remote sensing studies have explored the potential of optical satellite (Foody and Cox, 1994; Houghton et al., 2001; Lu et al., 2004), LIDAR data (Lefsky et al., 2002), or radar data (Austin et al., 2003; Rauste, 2005) for estimating forest biomass. Those studies as well as other references on forest biomass, however, mostly addressed tropical rainforests of the Amazon (Houghton et al., 2001).

Tab. 2.1 shows the state of research on biomass estimation in the tropics which is relevant to the present work. As mentioned before, these studies mostly focused on the regenerating- and secondary forests in tropical forest of the Amazon. The applications of moderate resolution of Landsat TM data are widely implemented in these studies. Although more accurate remote sensing data (e.g. Laser scanner (Lidar) or polarimetric Synthetic Aperture Radar (SAR) data) may improve the accuracy of biomass estimation of the forest regions, low temporal resolution and limited coverage of the data, which usually cover only very small portions of test sites, have restricted broader applications of these data, particularly in the tropical forest landscape.

2.4 Definition of Biomass

Biomass is generally defined as organic materials produced by plants, such as leaves, roots, seeds, and stalks. In some cases, microbes and animals are also included as biomass (Ashton, 2006). Ecology studies, however, have described biomass as the mass of living biological organisms in a given habitat or ecosystem at a given time.

Biomass can refer to species biomass, which is the mass of one or more species, or to community biomass, which is the mass of all species in the community. It can also include microorganisms, plants or animals (McNaught and A.Wilkinson, 1997). In this study, biomass is defined in a more specific manner, which is the total amount of above ground living organic matter in trees, expressed as oven-dry weight per unit area. This definition also refers to biomass density when expressed as mass per unit area, e.g. Mega gram (Mg)/ha (Brown, 1997).

Tab. 2.1 Recent studies on forest biomass assessment

AGB Estimate	Study Area	Forest Type	Climate/Topography	Methodology/Instruments	Data Used	Reference
Altamira: 140±130 Mg.ha ⁻¹ Bragantina: 100±90 Mg.ha ⁻¹ Pedras: 120±110 Mg.ha ⁻¹	Eastern Brazilian Amazon: Altamira, Pedras, and Bragantina	Secondary forest – different successional vegetation stages	Rainfall 2000-3000 mm Mixture of flat and rugged terrains. The elevation ranges from 50 – 300 m	Remote Sensing (RS) approach – TM bands, vegetation indices, band ratios, image transform (e.g. Principal component analysis, Tasseled cap)	Landsat TM	(Lu et al., 2004)
285±57 Mg.ha ⁻¹	Rondonia State, Southwestern Brazilian Amazon	Moist forest – dominated with open forest, some palms and lianas	Annual rainfall 2300 mm Topography: flat to undulating Elevation: 80 – 140 m	Allometric equation based on destructive sampling approach	Field measurement data	(Brown et al., 1995)
Brazil: 72 and 76 Mg.ha ⁻¹ Bolivia: ranged from 24 – 134 Mg.ha ⁻¹	Bolivia and Brazil	Regenerating secondary forest	-	RS approach – TM band 3,4,5 validated with allometric equations	Landsat TM	(Steininger, 2000)
Estimate: 225 – 486 Mg.ha ⁻¹ SD: 62 – 202 Mg.ha ⁻¹	Rondonia State, Southwestern Brazil	Open moist tropical forest	Annual rainfall 2600 mm Topography: flat to undulating Elevation: 80 – 140 m	Stem volume – AGB equation and Kriging method	Field data (RADAMBRASIL database)	(Sales et al., 2007)
20 – 160 Mg.ha ⁻¹	Luquillo Mountains, Puerto Rico	Mountainous tropical forest - early successional Forest	-	RS approach – Normalized Difference Vegetation Index (NDVI)	Landsat MSS and TM, Thematic Mapper Simulator, airborne multispectral scanner	(Sader et al., 1989)
Regenerating forest: 15 – 157 Mg.ha ⁻¹ Mature forest: 387 Mg.ha ⁻¹	Manaus and Tapajos forests, Brazil	Mainly regenerating forests with some mature forest patches and pasture	-	RS approach – radar backscatter (σ) and GLCM texture based allometric equations	JERS-1 SAR image with L band	(Kuplich et al., 2005)
Moisture gradient regenerating forests: 4 – 310 Mg.ha ⁻¹	Hawaii Volcanoes National Park	Mixture of two tiered wet forest, single tier mesic forest, open woodland and very scattered woodland	Rainfall 1270 – 2540 mm	RS approach – multi polarization (HH, HV, VV) radar backscatter (σ) and polynomial regression model	NASA/JPL Airsar data with C, L, and P band	(Imhoff, 1995)
Brazil: 76 – 421 Mg.ha ⁻¹ Thailand: 29 – 329 Mg.ha ⁻¹ Malaysia: 65 – 639 Mg.ha ⁻¹	Manaus (Brazil), Danum Valey (Malaysia) and Khun Khong (Thailand)	Brazil: dense tropical forest, Malaysia and Thailand: logged over forests	-	RS approach – vegetation indices, complex band ratios complemented with multi-linear regression and neural networks method	Landsat TM	(Foody et al., 2003)
Sungai Wain: up to 400 Mg.ha ⁻¹ Mawas: 20 – 350 Mg.ha ⁻¹	Mawas and Sungai Wain, Kalimantan, Indonesia	Tropical lowland dipterocarp and peat swamp forests	Annual rainfall: >2000 mm Topography: flat to undulating	RS approach – Random Volume over Ground (RVoG) model and inversion of dual-polarization model	Airborne multi-band (C, L, P, X band) and multi-polarization (Polinsar) data	(Hajnsek et al., 2009)

The terms of dry weight biomass is used here to distinguish from the fresh weight biomass prior to removal of water contents. while the term of above ground biomass (AGB) is applied to differentiate from below ground biomass (BGB) that contains of all living roots over 2 mm in diameter (Ravindranath and Ostwald, 2007).

The AGB of many ecosystems accounts for about 70 – 80% of total biomass and acts as an important carbon pool for many vegetation types and land use systems (Ravindranath and Ostwald, 2007). Replanting of degraded land leads to continual accumulation of AGB whereas any disturbance to vegetated areas leads to loss of the AGB. Therefore, it is important to study this pool for most carbon mitigation as well as other land-based projects. Estimation of stock changes in the AGB is also necessary for greenhouse gas inventory at national level for different land use categories such as forest lands, cropland and grassland. This present study presents methods of estimating and monitoring above ground biomass from remotely sensed data and GIS combined with field measurement data by developing an alternative, non-destructive approach of AGB estimation.

2.5 Retrieval of Forest Biomass using RS Data

Retrieval of forest stand properties, such as biomass, from remote sensing data means that the forest properties are modeled using spectral information from the satellite data, and the predicted results are then validated using a set of field observation data. The applicability of remote sensing data for environmental studies is growing rapidly since the issues of global warming and accelerated carbon emission rate have become of major concern around the globe (UNFCCC, 1992).

Forest biomass assessments have greatly benefited owing to the availability of optical sensors, especially Landsat data (Tab. 2.1). The wide application of this particular sensor is due to its spatial and temporal resolution and wide coverage area. Application of the Landsat data for biomass estimation is relatively straightforward and the model complexity mostly depends on the selected regression model. Additionally, ancillary data (e.g. vegetation indices, simple ratios, image transforms, etc) to increase the estimation accuracy can be easily generated from the multi-spectral bands and fused together with the spectral information (Lu et al., 2004).

Vegetation indices generated from the multispectral bands are sensitive to characterize green vegetation/forested regions from other objects on the ground (Jensen, 1996). In vegetated regions, the cells in plant leaves are very effective scatterers of light because of the high contrast in the index of refraction between the water-rich cell contents and the intercellular air spaces. Vegetation is very dark in the visible bands (400-700 nm) because

of the high absorption of pigments which occur in leaves (chlorophyll, protochlorophyll, xanthophyll, etc.). There is a slight increase in reflectivity around 550 nm (visible green band) because the pigments are least absorptive. In the spectral range of 700-1300 nm plants are very bright because this is the spectral range of electronic transitions, providing absorption in the visible and molecular vibrations that absorb in longer wavelengths. There is no strong absorption in this spectral range, but the plant scatters strongly. From 1300 nm to about 2500 nm vegetation is relatively dark, primarily because of the absorption by leaf water. Cellulose, lignin, and other plant materials are also absorbed in this spectral range (Lillesand and Kiefer, 1994).

The main problem of modeling forest biomass using Landsat data is the saturation of vegetation indices at higher biomass density, as reported in previous studies (Lu, 2005; Steininger, 2000). Various factors might cause the saturation of the Landsat data, e.g. high variability and complex structure of forest attributes and limited range of vegetation indices (Lu, 2006). Evaluation of input data variations for modeling the biomass, combining spectral data, vegetation indices, simple ratios, and image transformation generated from the Landsat data, digital elevation model (DEM), and field observation data are explored in this work. We propose to model above ground biomass (AGB) using a non-destructive approach regardless of destructive tree sample collection.

On the other side, microwave remote sensing allows the provision of additional information measured from the ground objects, since they are insensitive to the cloud-free daylight conditions for image acquisition. Application of SAR data for mapping of forest properties has already been widely applied (Fransson and Israelsson, 1999; Hajnsek et al., 2005; Isola and Cloude, 2001; Kuplich et al., 2005; Luckman et al., 1996). Besides, empirical models of microwave instrument data are known to be very sensitive to the density, shape, length, dielectric properties, and orientation of the scatterers (Kingsley and Quegan, 1992). The X-band (2.4 – 3.75 cm) SAR data is useful for terrain mapping and for discriminating the top canopy of vegetated land. Some studies showed that the utility of single polarization C-band data (3.75 – 7.5 cm) may bring some limitations for distinguishing biomass in regenerating forests and deforested areas, because the radar backscatter becomes insensitive especially if the soil is dry and the influence of water is minimized (Saatchi et al., 1997).

In contrast, L-band SAR data (15 – 30cm) showed good ability for modeling the forest parameters under dense vegetation (Luckman et al., 1997; Rauste, 2005). The capability of L-band radar backscatter to penetrate through the forest canopy makes this data useful for mapping the forest structure, including above ground biomass (AGB) estimation (Luckman

et al., 1997). Recent studies confirmed that the empirical AGB model generated from the L-band radar backscatter provides favorable estimation results, especially if the data has dual-polarization configuration. Using dual polarimetry SAR data the backscatter signal is more sensitive to forest biomass and forest structure because of tree trunk scattering thus showing some discrimination between different forest succession stages (Ranson et al., 1997; Rignot et al., 1997; Saatchi et al., 1997). Similar to vegetation indices generated from optical RS data, the biomass estimation from radar backscattering is also limited up to a certain biomass level due to saturation problems. The backscatter saturation limits are dependent on the geometry of data acquisition, the polarization compositions, the wavelength, and the complexity of vegetation structure. These values may vary from 20 – 40 Mg/ha at C-band to close to 100 Mg/ha at L-band (Imhoff, 1995). However, in very dense tropical rain forests it is common to find biomass over 300 Mg/ha which is far beyond the suggested saturation limit.

The use of polarimetry SAR interferometry (Polinsar) data is an alternative approach to estimate AGB from stem volume more accurately. The coherence of interferometry SAR data is used to generate a digital surface model, which is used as a basis to estimate tree height. Using random volume over ground inversion model the tree volume is estimated, thus above ground biomass is predicted using allometric models (Hajnsek et al., 2009; Mette et al., 2003). The estimation of biomass using this approach could eliminate the saturation problem, but to accurately model tree height from coherence values is somehow problematic, since to optimize the interferometry coherence one has to deal with the minimization of the baseline and temporal decorrelation of the image pairs.

Some studies have mentioned that the individual use of SAR data does not provide complete information for land cover classification and the extraction of forest biophysical parameters, and additional information can be provided if a microwave sensor is combined with optical data (Kuplich, 2006; Rauste, 2005; Rignot et al., 1997). These authors found that fusion of both sensors increased the discrimination of different forest regenerating stages. The classification procedures combining microwave and optical sensors are yet under development and remain a great challenge for remote sensing studies (Lu, 2006). Here, we are experimenting with the combination of both sensors for characterizing different classes on tropical forests landscape and peatland forests.

2.6 Biomass Model Transferability and Issues on the Model Accuracy

Foody (2003) found that the transferability of biomass models over different forest regions are limited not only due to forest structure and vegetation composition differences, but also

uncertainties of remotely sensed data that can arise from data handling and preprocessing. The issues on model transferability also occur for the application of multi-temporal datasets over the same forest regions, caused by different spectral data characteristics and disturbance from haze and atmospheric aerosols. Lu (2006) mentioned that the accuracy of estimated AGB in boreal forests was around 9%. However, Sales et al. (2007) obtained around 30 - 40% of error estimate for the AGB estimated in the tropical Amazonian forests. The uncertainty on biomass modeling and the issues of model transferability have encouraged us to conduct the present study to assess the proper knowledge on the biomass modeling and to evaluate the biomass estimation over complex vegetated landscape of tropical humid forests. The combination of optical sensor (i.e. Landsat data), texture features and vegetation indices for AGB estimation is considered in this study.

Chapter 3

Study of Geostatistics Texture Features for Complex Land Cover Mapping^{*}

ABSTRACT

Provision of accurate information from land cover maps derived using remote sensing data can effectively support sustainable forest management. Land cover maps were mostly classified using per pixel or spectral data. This approach ignored the potential of useful spatial information from proximate pixels. Although the benefits of spatial information have been greatly explored, there have been limited attempts to enhance image classification combining spectral and spatial information. This study aims to explore the potential of spatial texture variability for the classification of complex land cover. Band 5 and 4 of Landsat ETM image were selected as primary and secondary predictors for the estimation of Geostatistics texture features. We used semivariogram, madogram, rodogram, pseudo-cross variogram and pseudo-cross madogram in calculating the texture. Another texture features generated from fractal dimension was also estimated. Support Vector Machine (SVM), Back-propagation Neural Network and Maximum Likelihood Classification (MLC) methods were applied to classify the spectral data of Landsat ETM image and the texture features. The accuracy assessment obtained that spatial information provided from the texture features improved the accuracy of spectral data classification. This study recommended rodogram, madogram and fractal dimension texture features as useful ancillary data for spectral data classification.

Keywords: geostatistics, texture features, land cover classification, SVM, MLC, neural network, fractal

^{*} This chapter is based on:

Wijaya, A., P.R. Marpu and R. Gloaguen (2008). Geostatistical Texture Classification of Tropical Rainforest in Indonesia. *Quality in Spatial Data Mining*. Editor: Stein, A., Shi, S. and Bijker, W. CRC Book Series, pp. 199 – 210.

3.1 Introduction

Mapping of land cover is an ultimate way to assess changes on forest cover and to study forest resource dynamics within a range of time. Forest encroachments have been hardly reduced due to excessive human exploitation of forest resources. In the last two decades, forest degradation and deforestation in tropical regions occur at an alarming rate. Tropical forests in Indonesia are also under serious threats, mainly due to anthropogenic factors. Regular monitoring on forest cover and its changes is important to support the objectives of sustainable forest management. The provision of accurate and updated land cover information over the forested lands may support these objectives. Accurate classification of satellite data can efficiently and rapidly monitor huge forest areas at relatively low cost. Interpretations of satellite image data usually take into account per pixel classification rather than correlation with its neighboring pixels. In this study, we propose geostatistics approach that considers spatial variability among neighboring pixels (Jakomulska and Clarke, 2001; Wijaya et al., 2007). Geostatistics and the theory of regionalized variables were introduced to remote sensing by Woodcock et al. (1988). The application of geostatistics texture feature for remote sensing study has been made popular by several studies (Miranda and Carr, 1994; Miranda et al., 1998; Miranda et al., 1996; Miranda et al., 1992).

This work contributes to provide alternative texture features estimated using multivariate geostatistics and fractal dimension. Another objective was to combine textures features and spectral data for satellite image classification comparing Maximum likelihood classification (MLC), Support Vector Machine (SVM) and Multi Layer Perceptron Neural Network (MLP-NN) methods. We also identified the texture features that were useful for the classification of tropical forests with complex vegetation structure. Texture in this work is defined as a representation of the digital number (DN) values variation in an satellite image, which provides important information about the structural arrangements of the image objects and their relationship to the environment (Chica-Olmo and Abarca-Hernandez, 2000). This work was greatly encouraged by the following factors: (1) texture features providing extra information should improve the classification results; (2) image classification over forest regions where no dominant distinct objects (i.e. objects with particular shape or boundary) were exist can benefit from the additional information provided by the texture variation; and (3) provided similar spectral values, the texture arrangement of these values was distinguishable so that the use of texture feature should improve the classification accuracy (Asner et al., 2002; Kayitakire et al., 2006; Wijaya et al., 2007).

Estimation of geostatistics texture information considers spatial correlation of a single pixel with respect to its neighboring pixels (Chica-Olmo and Abarca-Hernandez, 2000). In the case of

supervised classification, texture estimates should take into account spatial variability of training data and the relation of which with their neighboring pixels. This means, in a satellite image, variations of similar adjacent DN values which represent certain types of land cover should be first recognized before texture data are estimated. The information on this spatial variation was used to select the size of moving window for calculating geostatistics texture data. Upon selection, geostatistics texture data were calculated using univariate spatial variability measures (i.e. semivariogram, madogram, rodogram, and fractal dimension) and multivariate measures (i.e. pseudo-cross semivariogram and pseudo-cross madogram).

3.2 Satellite Image Preprocessing

3.2.1 Atmospheric correction approaches

This study used a 30 m resolution Landsat 7 ETM+ image acquired on May 31, 2003. The image was geometrically corrected using 30 ground control points and resampled using nearest neighborhood algorithm to minimize image histogram changes. The root mean square error was maintained below one pixel resulting in an accurate orthorectified image. For atmospheric corrections of the ETM data several approaches are available. Some are relatively straightforward while others are complex, being founded on physical principles and requiring a significant amount of information to function properly (Jensen, 1996). In general, there are two types of atmospheric corrections, which are absolute and relative corrections. Absolute atmospheric correction converts digital number (DN) values into surface reflectance, while relative correction converts DN values of multi temporal satellite images to the same reflectance regardless the actual condition may be on the ground (Song et al., 2001).

Absolute correction of the image data for the effects of atmospheric propagation can be carried out in essentially three ways. Physically based methods attempt to model the atmospheric scattering and absorption on image data based on computer modeling, or so-called Radiative Transfer Modeling. Two best known methods applying the physical based correction are LOWTRAN and MODTRAN (Cooley et al., 2002). The second approach to atmospheric correction of satellite imagery is based on calibration against targets of known reflectance (Rees, 2001). These targets can be artificially constructed or naturally occurring, but their reflectance should be accurate and sufficient, and the availability of well-distributed sensor calibrators within the corrected image scene is necessary. Finally, the simplest, and most widely applied method of atmospheric correction is based on dark pixel subtraction or dark object subtraction (DOS) (Chavez Jr., 1996). This approach assumes that in every image scene dark objects are exist, thus the radiometric correction can be done by

subtracting the pixel values of dark object from all pixels in the image scene. Using DOS technique only the atmospheric effects caused by atmospheric scattering is corrected, because this method does not model a complex atmospheric absorption effect in its equation (Lu et al., 2004). Another assumption that should exist is that the earth surface should be a Lambertian surface, i.e. a surface which reflects radiated radiance equally in all directions.

The selection of atmospheric and radiometric corrections to the image data depends upon the application and the accuracy of corrected reflectance needed for the purpose of the study. While physically based method may be the most accurate means of correcting for atmospheric effects such a method requires information on atmospheric properties that are typically unavailable. Therefore, simple methods are often quite effective for atmospheric correction of remote sensing data (Foody et al., 2003; Schroeder et al., 2006; Song et al., 2001).

3.2.2 Radiometric calibration and topographic corrections methods

Radiometric calibration is a multi-step process that involves the use of standard equations to convert 8-bit satellite-quantized calibrated digital numbers (DN) to at-satellite reflectance. Landsat 7 images were converted to at-satellite radiance using Eq. 3.1,

$$L_{sat} = ((Lmax_{sat} - Lmin_{sat}) / (DNmax - DNmin)) \times (DN - DNmin) + Lmin_{sat} \quad \text{Eq. 3.1}$$

where $Lmax_{sat}$ is band-specific spectral radiance scaled to DNmax ($W\ m^{-2}\ sr^{-1}\ \mu m^{-1}$), $Lmin_{sat}$ is band-specific spectral radiance scaled to DNmin ($W\ m^{-2}\ sr^{-1}\ \mu m^{-1}$), DNmax is maximum quantized calibrated digital number (255), and DNmin is minimum quantized calibrated digital number (0 for LPGS data, 1 for NLAPS and EOS data). Equation 3.1 accounts for gain state (i.e. high/low settings) by using respective published $Lmax$ and $Lmin$ values as published in Landsat 7 Science Data User Handbook (Landsat Project Science Office, 1998).

Converting to at-satellite radiance, each Landsat band was converted into top-of-atmosphere reflectance (ρ_{TOA}) using following equation (Lu et al., 2004),

$$\rho_{TOA} = \frac{\pi \cdot d^2 \cdot L_{sat}}{E_{sun} \cdot \cos(\theta)} \quad \text{Eq. 3.2}$$

where d^2 is square of the earth-sun distance in astronomical units (Au) that can be found in Landsat 7 Science Data User Handbook (Landsat Project Science Office, 1998), E_{sun} is mean solar exo-atmospheric irradiance ($W\ m^{-2}\ \mu m^{-1}$), and θ is sun zenith angle.

The conversion from at-satellite radiance to top of atmosphere reflectance is known as partial correction case since by definition this value does not remove atmospheric effects due to atmospheric scattering.

For absolute atmospheric correction case, the at-satellite radiance is converted to surface reflectance (ρ), assuming a uniform Lambertian surface under cloudless condition, using following formula (Chavez Jr., 1996; Schroeder et al., 2006),

$$\rho = \frac{\pi \cdot d^2 \cdot (L_{sat} - L_p)}{T_v \cdot (E_{sun} \cdot \cos(\theta) \cdot T_z + E_{down})} \quad \text{Eq. 3.3}$$

where L_p is path radiance ($W\ m^{-2}\ sr^{-1}\ \mu m^{-1}$) or so-called haze layer (Chavez Jr., 1988), T_v is atmospheric transmittance from the target toward the sensor, T_z is atmospheric transmittance in the illumination direction, and E_{down} is downwelling diffuse irradiance ($W\ m^{-2}\ \mu m^{-1}$). The output of atmospheric correction is the percentage of surface reflectance. There are different variations of DOS technique according to a study conducted by Song et al. (2001). This study initially adopted two variations of DOS technique, which are DOS1 and DOS2. We found afterward that DOS2 has better correction result in terms of the spectral response of corrected image in relation with the actual spectral of vegetation types. Thus, hereafter the satellite image was corrected using DOS2 technique.

In flat terrain condition, illumination effects due to local zenith angle is minor, but in rugged terrain areas due to different topographic position this effects is noteworthy and thus topographic effects should be corrected even if the areas are directly illuminated by the sun (Mather, 2004). Topographic effect causes a high variation in the reflectance response for similar vegetation types; shaded areas show less than expected reflectance, whereas in sunny areas the effect is the opposite (Riaño et al., 2003).

One way to correct topographic effects is by assigning the Digital Elevation Model (DEM) of the area. The DEM should be firstly geo-referenced in the same coordinate system as the satellite image to be corrected and the DEM should also be of a scale that is close to that of the satellite image, so that the slope angle and aspect can be derived for each pixel position of the satellite image. Thus, the DEM is also used to compute the incident angle (γ_i), defined

as the angle between the normal to the ground and the sun rays (Riaño et al., 2003). The γ_i parameters varies from -1.0 to +1.0, and are computed as:

$$\cos \gamma_i = \cos \theta_p \cos \theta_z + \sin \theta_p \sin \theta_z \cos (\phi_a - \phi_o) \quad \text{Eq. 3.4}$$

where θ_p is the slope angle; θ_z is the solar zenith angle; ϕ_a is the solar azimuth angle; and ϕ_o is the aspect angle; each angle should firstly be converted into radian.

Thus, the incident angle was estimated for the whole image and the topographic correction can be carried out using available methods. A simple method to correct terrain slope in areas that receive direct solar illumination is simply to adopt Lambertian assumption (i.e. a surface reflects radiation in a diffuse fashion, so that it appears equally bright from all feasible observation angles). Instead of applying terrain correction with Lambertian assumption, this study is more cautious of diffuse irradiance since most rugged terrains basically have a non-Lambertian behavior. This non-Lambertian correction models include Minnaert correction (Minnaert, 1941), which is probably the most popular empirical method for computing a complicated Bidirectional Reflectance Distribution Function (BRDF) of rugged terrain areas. This study however implemented another empirical-statistical method assuming a linear correlation between the reflectance of each satellite band and the incident angle cosine ($\cos \gamma_i$) as (Teillet et al., 1982)

$$\rho_T = b_k + m_k \cdot \cos \gamma_i \quad \text{Eq. 3.5}$$

where m_k is the slope of the regression line for band k . The b_k is considered constant for the entire image, being the intercept in the regression equation. A variation of this empirical approach is called the C-correction method (Teillet et al., 1982), and the correction of the reflectance under an inclined surface (ρ_T) is defined as

$$\rho_H = \rho_T \left[\frac{\cos \theta_z + (b_k/m_k)}{\cos \gamma_i + (b_k/m_k)} \right] \quad \text{Eq. 3.6}$$

where θ_z is the solar zenith angle, and ρ_H is the reflectance of a horizontal surface.

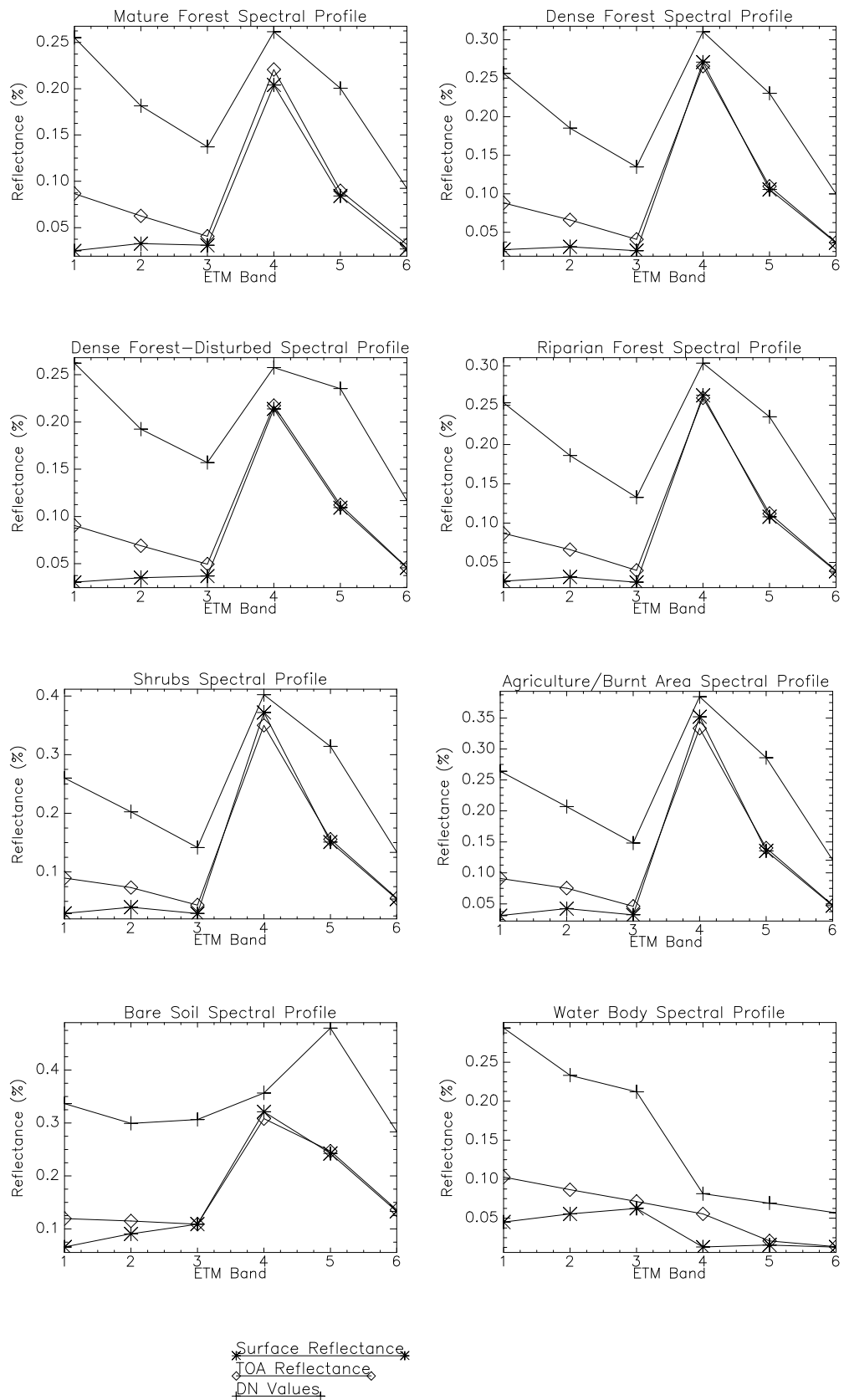


Fig. 3.1 Spectral responses of multispectral bands Landsat ETM data for different land cover types (band 6 in the plot represents far infra red band of ETM band 7). The relative reflectance values clearly show the effects of atmospheric attenuations in partially corrected image (top of atmosphere reflectance profiles) as compared to surface reflectance profiles

3.2.3 Spectral profiles and histogram of the Landsat ETM image

We collected spectral sample of calibrated image, partially corrected image (i.e. top of atmosphere (TOA) reflectance), and uncorrected images from different forest physiognomies, namely mature forests, very dense forests, dense forests, riparian forests and shrubs, as presented in Fig. 3.1.

The band 6 in the horizontal axis of Fig. 3.1 refers to the ETM band 7, as we plotted only the ETM multispectral bands (ETM band 1 – 5, and band 7). Comparing those spectral responses, we obtained that the DN value was simply a number scaled from 0 – 255 showing the brightness of different objects on the earth surface, therefore, it could not be directly compared, both with TOA reflectance and surface reflectance. The spectral profiles of atmospherically corrected image in all forested lands and bare soil classes expressed similar patterns, which were lower values at visual bands (ETM band 1, 2, 3) and higher values at infrared bands (ETM band 4, 5, and 7). The spectral profile of each land cover type shows that the partially corrected image is yet affected by atmospheric attenuations. These effects, which are more prominent for visual bands, are mainly caused by atmospheric scattering and haze. The atmospheric corrections using DOS technique greatly reduced the attenuation effects and yielded more favorable image for remote sensing analysis. The infrared bands (ETM band 4, 5, and 7) on the other hand, are less sensitive to the atmospheric effects because the reflectance values of partially corrected and atmospherically corrected images were slightly different. Non-vegetated areas, i.e. bare soils and water body, are less sensitive to the atmospheric effects, yet the radiometric calibration improved the spectral profiles of both land cover classes.

The histograms of uncorrected ETM data, the atmospheric corrections using DOS and MODTRAN methods were calculated and compared (Fig. 3.2). The DN values of the uncorrected image confirmed that ETM band 1, 2 and 3 mostly suffered from the atmospheric attenuations. Physical-based atmospheric corrections based on MODTRAN method, normalized the surface reflectance from 0 – 10,000 (for 0.0 to 1.0 interval), corrected these effects accurately. This method considered not only the atmospheric scattering and haze in the corrections, but also the atmospheric absorption. Unfortunately, the result could not be validated, due to the absence of measured spectral values on the ground. The histograms of MODTRAN method and DOS technique were compared, and each ETM band showed respectively similar data distribution, only the reflectance of the earlier method were about two times of the latter. Limitations of DOS approach for correcting the atmospheric scattering and to accurately model the aerosol type and visibility may be the underlying factors of this difference.

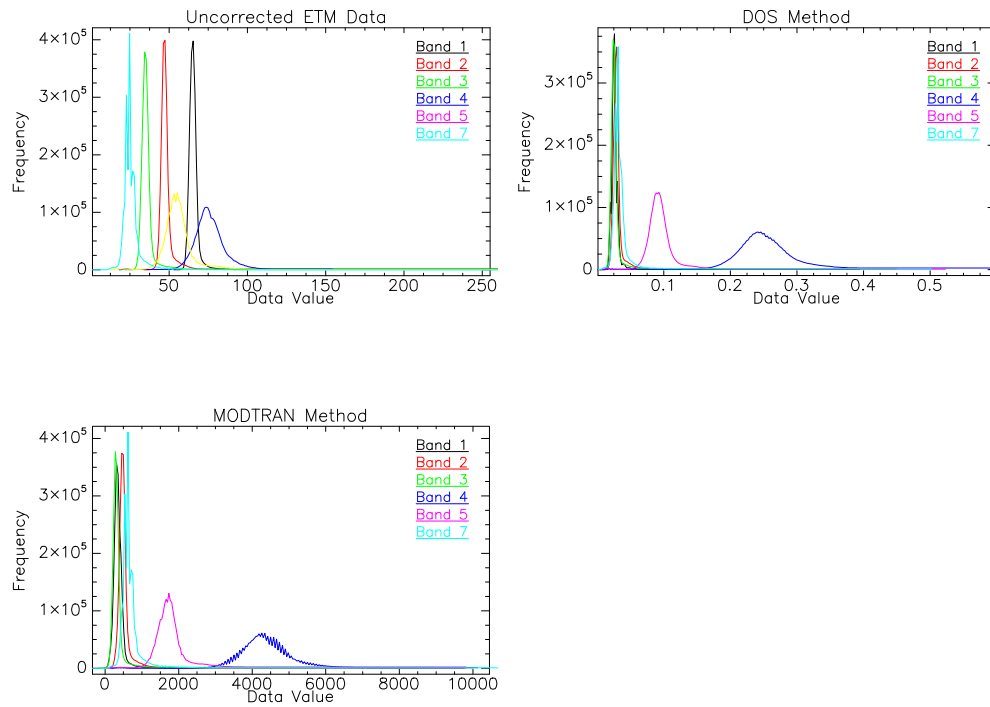
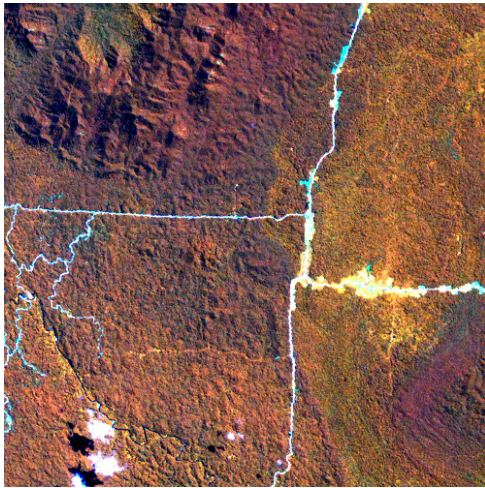


Fig. 3.2 Histogram comparison of the uncorrected ETM image and corrected images based on a physical based MODTRAN method, and an image based DOS corrections approach. Both methods greatly reduced the attenuation effects in visual and near infra red bands and the histograms of the ETM image, either corrected using MODTRAN or DOS methods, show similar data distribution.

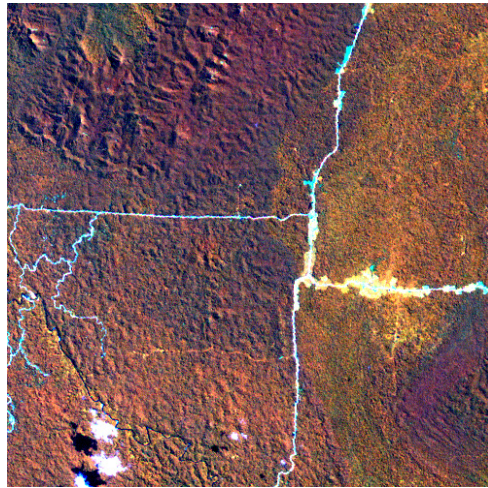
3.2.4 Image corrections results

The C-correction method is a simple method which is used to reduce the shadows and topographic effects very efficiently. This method is easy to implement and could provide an adequate result that is sufficient for the purpose of this study. In topographically corrected images, the shadows due to terrain slope were minimized, although the terrain effects were not extreme, as the study site was actually dominated by flat terrains (Fig. 3.3). Another example taken from an image subset representing a mixture of complex vegetation and rugged terrains also delivers a positive result showing reduced terrain effects upon the corrections (Fig. 3.3c and Fig. 3.3d). Successful topographic corrections in the preprocessing stage may improve interpretation and classification of the satellite image.

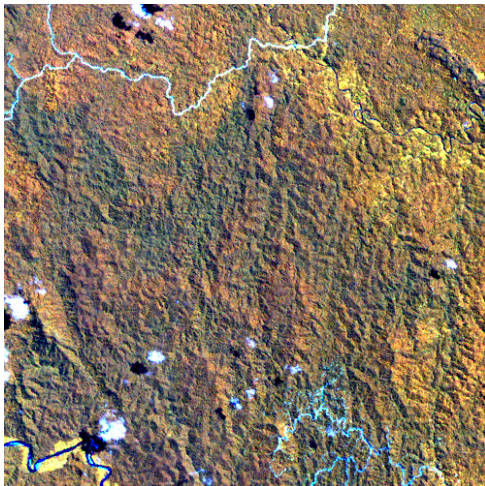
Compared to ATCOR3 module—an integrated atmospheric and topographic corrections module which is commercially available in PCI Geomatics software with extra license, the C-correction method, whose script is written in IDL programming language, is available for free in the internet. Hereinafter, the Landsat ETM image used for the entire study is the resulted products of the atmospheric and topographic corrections presented in this chapter.



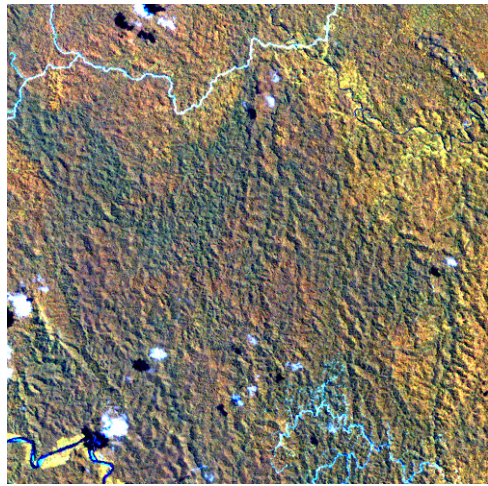
a) Uncorrected image



b) Topographically corrected



c) Uncorrected image



d) Topographically corrected

Fig. 3.3 The result of topographic corrections in two subsets of the study site. The topography effects that cause shadows in the back side of the slope were reduced, although the study area is located in relatively flat region.

3.3 Geostatistics texture features

Texture layers using geostatistics can be computed over each band of the satellite data. However, not all the texture features are useful for spectral data classification. Therefore, pre-selection of the satellite band containing most texture information should be conducted. For each satellite band we computed the covariance matrix showing the variance of each land cover class, and two satellite bands correspond to the highest mean variance of forest classes were selected, because we need primary and secondary predictors for estimating multivariate semivariance.

To incorporate geostatistical texture features in spectral classification, we computed the semivariogram from each neighboring pixel. Generally, spatial variability $\gamma(h)$ increases gradually with distance separating the observations up to a maximum value (the sill) representing the maximum spatial variance. The distance at which the sill is reached represents the range of variation, i.e., the distance within which observations are spatially dependent. Respectively, the size of moving window used to extract texture information of spectral data has an important role in providing an accurate estimation of semivariance, which eventually affects the classification accuracy.

Semivariogram is an univariate estimator, which describes the relationship between similarity and distance in the pixel neighborhood. $Z(x)$ and $Z(x + h)$ are two values of the variable Z located at points x and $x + h$. The two locations are separated by the lag, h . The semivariogram values, $\gamma(h)$ are calculated as the mean sum of squares of all differences between pairs of values with a given distance divided by two as described in the following equation (Carr, 1995).

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n (Z(x_i) - Z(x_i + h))^2 \quad \text{Eq. 3.7}$$

Where n is the number of data pairs.

The madogram, instead of measuring squares of all differences takes the absolute values. Madogram is simply a first-order variogram and has its roots in traditional geostatistics. Avoiding the squared term, madogram has a more robust measure of the spatial correlation structure than semivariogram and it has a convenient relationship with multi-variate extreme value. Madogram was calculated using following equation (Chica-Olmo and Abarca-Hernandez, 2000; Deutsch and Journel, 1998).

$$\gamma_m(h) = \frac{1}{2n} \sum_{i=1}^n |Z(x_i) - Z(x_i + h)| \quad \text{Eq. 3.8}$$

By calculating square root of absolute differences, we can derive a spatial variability measure which can reduce the impact due to the presence of data outliers, called rodogram. The rodogram was estimated as follow (Lloyd et al., 2004).

$$\gamma_r(h) = \frac{1}{2n} \sum_{i=1}^n |Z(x_i) - Z(x_i + h)|^{\frac{1}{2}} \quad \text{Eq. 3.9}$$

Spectral data of satellite image mostly have correlation effect particularly among two respective bands (Wijaya et al., 2008b). Given the fact that this correlation effect can provide useful texture information, we computed the multivariate estimators quantifying the joint

spatial variability (cross correlation) between two bands, namely pseudo cross variogram and pseudo cross madogram. The pseudo-cross variogram represents the semivariance of the cross increments calculated as follows:

$$\gamma_{PCV}(h) = \frac{1}{2n} \sum_{i=1}^n (Y(x_i) - Z(x_i + h))^2 \quad \text{Eq. 3.10}$$

The pseudo-cross madogram is similar to the pseudo-cross variogram, but again, instead of squaring the differences, the absolute values of the differences area were taken, which leads to a more generous behavior toward outliers (Buddenbaum et al., 2005).

$$\gamma_{PCR}(h) = \frac{1}{2n} \sum_{i=1}^n |Y(x_i) - Z(x_i + h)| \quad \text{Eq. 3.11}$$

The fractal dimension was used in various applications (Biswasa et al., 1998; Chaudhuri and Sarkar, 1995; Jaenisch et al., 1993; Rogers et al., 1995). In pattern recognition, the fractal dimension is defined as a measure for the coarseness of objects and textures, which can be assigned to a gray level image to get the texture measure of the image. In remote sensing, fractal dimension is considered as a good texture measure that can improve classification accuracy (Chica-Olmo and Abarca-Hernandez, 2000). For texture segmentation, the fractal dimension can be calculated using (Gloaguen et al., 2007)

$$\gamma(h) \propto 2|h|^H \quad \text{Eq. 3.12}$$

where $\gamma(h)$ is the semivariance at lag h and H is defined as the slope $\log \gamma(h)$ vs $\log h$. Given the semivariogram of any spatial distribution, fractal dimension (D) of the area is commonly estimated as follows (Burrough, 1983; Carr, 1995; Gloaguen et al., 2007).

$$D = 2 - \frac{H}{2} \quad \text{Eq. 3.13}$$

Fractal dimension is very effective in improving classification accuracy providing useful texture information image regions (Wijaya et al., 2007).

To observe textural variation among different land cover classes which corresponds to lag distance, semivariance values of training dataset for each land cover class were sequentially computed using a 7×7 window for up to 8 neighboring pixels (lags). The computed semivariance provide useful information for data classification as these values revealed spatial correlation of each land cover type with respect to the lag distance.

The estimation of texture features was explained as follows. For example, in a 5×5 window size, every adjacent pixel in the NSEW direction was considered as semivariance at 1, 2, 3 and 4 lags. For diagonal direction (e.g. $\sqrt{2}$, $\sqrt{3}$, $\sqrt{4}$, etc) the lags were rounded up or down to

the closest integer value. Using Eq 3.7 – 3.13, median of semivariance values in every lag was calculated and used as a new DN value for the respective pixel (see appendices for the scripts used for generating the geostatistics texture features).

3.4 Classification methods

3.4.1 Support Vector Machine Classification

The SVM is a type of universal learning machine, which is used for pattern recognition and was originally designed to solve binary classification problems. Details of the SVM principle may be found in several recent publications (Bazi and Melgani, 2006; Foody and Mathur, 2006; Scholkopf and Smola, 2001; Vapnik, 2000). Two main approaches have been developed for multi-class classification using SVM, i.e. the One-Against-All (OAA) and the One-Against-One (OAO) methods (Hsu and Lin, 2002). The OAA method compares one class with all the others taken together. For n classes, n hyper planes are determined, n optimization problems need to be solved and n classifiers are generated. The OAO approach performs a binary SVM on all possible pairs out of n classes. Each classifier is trained on two out of n classes, and number of classifiers therefore is $n(n - 1)/2$. Applying these classifiers to a test data point leads to $n(n - 1)/2$ class votes (Hahn et al., 2007). The test data is then labeled from the class that received the most votes. In general, the OAA and OAO reduce the multi-class dataset into several binary problems that have to be solved (Hsu and Lin, 2002). This study subjectively implemented multi-class OAO SVM classification using LIBSVM software developed by Chang and Lin (2001).

Multi-class image classification using the SVM method is conducted by combining several binary classifications with the support of optimum hyperplane. The optimum performance of this method is mainly affected by a proper selection of kernel parameters involved in the algorithm (Wijaya and Gloaguen, 2007a). For the classification, Radial Basis Function kernel was used. Estimation of the generalized accuracy using different kernel parameters (γ) and cost parameters (C) were applied with the following intervals: $\gamma, C = \{2^{-15}, 2^{-13}, 2^{-11}, \dots, 2^7\}$. For each possible kernel parameters, SVM classifications were carried out and its accuracies were assessed using confusion matrices. We divided the training dataset into five subsets and each time one subset was used for the classification and its accuracy was assessed using other subsets. The iterations were repeated for other subsets, and the input data, kernel parameter (γ) and cost parameter (C) which yielded highest accuracy were recorded.

First of all the classification iterations were performed using each possible combination of three out of six ETM multispectral bands, and then continued with four and five ETM bands iterations. We assessed the classification accuracies using confusion matrices, and recorded

the input variations of three, four and five ETM bands with highest accuracies, as well as their kernel parameters.

3.4.2 Multi Layer Perceptron Neural Network

Neural Network are essentially learning systems based on interconnected networks of simple processing elements (Wijaya, 2006). In general, there are three phases in neural network classification. The first phase is a training procedure using input data. The second is a validation phase that determines the success of training phase and the accuracy of the network when it is applied to predict unseen data. The last stage is classification of a full image (Gahegan et al., 1999). The use of neural network for multi source data classification was made popular because the data distribution needs no normality assumption. Due to different data distributions, the combination of spectral and texture data could violate normality assumption that was required by most of statistical classifiers, e.g. maximum likelihood (Tso and Mather, 2001). This study implemented three-layered neural network consisting of a single input, hidden, and output layer, so-called multi-layer perceptron neural network (MLP-NN), and the network was trained by back-propagation algorithm. The MLP neural network trained by back-propagation algorithm was commonly used for image classification in remote sensing study (Kanellopoulos and Wilkinson, 1997). Training iterations were carried out by estimating root mean square error (RMSE) of gradient descent. The training was stopped, either when total RMSE was less than 0.0001 or before the cross-validation accuracy of training data started decreasing. Sigmoid model was selected as the transfer function of the neural network.

3.4.3 Variations of Input Data

Given the cross validation results from SVM classification, we found that the combinations of ETM band 3,4,5; ETM band 1,2,4,7; and ETM band 1,2,3,4,7 outperformed other combinations. Our experiments found that band combinations 1,3,4,5 and 1,2,3,4,5,7 (multi-spectral bands) of ETM data have similar accuracies with those of earlier mentioned spectral data. Therefore, further classification considered these ETM data variations.

We applied the same approach to find combinations of texture features which are more useful in improving classification accuracy. We obtained that the two geostatistics texture features, fractal dimension and median of rodogram outperformed other texture variations. Different result was found when three texture layers were introduced in the SVM classification as fractal dimension, variogram and pseudo-cross variogram yielded better accuracy compared to other texture features. Subsequently, four texture features were used as additional information and the accuracy assessment showed that fractal dimension,

rodogram, madogram and pseudo-cross variogram have better accuracy compared to other texture variations. The experiments with Grey Level Co-occurrence Matrix (GLCM) texture features were conducted by estimating second order statistics, namely variance, contrast, homogeneity, and dissimilarity (Haralick et al., 1973). However, classification using geostatistics texture features provided higher accuracy than of using GLCM texture features. Thus, we did not consider GLCM texture features for further analysis.

3.5 Results and Discussion

3.5.1 Texture features generation

For forest classes the middle infra red band (ETM Band 5) and near infra red band (ETM Band 4) have higher mean variance compared to ETM Band 1,2,3 and 7 (Tab. 3.1). Thus, univariate geostatistical texture features, i.e. fractal, variogram, rodogram, and madogram, were estimated using ETM band 5. Computation of multi-variate texture features (i.e. pseudo-cross variogram and pseudo-cross madogram) which needs two predictors were performed using ETM band 5 and ETM band 4 as primary and secondary predictors, respectively. Logged over forest, burnt areas/open forest, dense forest and hill shadow classes were spatially correlated within 5 lags, whilst road network and clear cut forest/bare land showed spatial correlation of more than 7 lags (Fig. 3.4). As we concerned for the classification of forest classes, a 5×5 window size was selected in calculating the texture features.

Tab. 3.1 Variance matrix of forest cover classes training data based on Landsat ETM image

Land Cover Class	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
Logged Over Forest	3.49	3.03	8.46	23.56	50.98	0.70	15.46
Burnt Areas/Open Forest	2.59	1.76	2.59	24.19	18.46	0.46	5.04
Road Network	65.68	165.39	332.63	74.12	386.45	1.91	294.98
Clear Cut Forest/Bare Land	2.70	5.20	2.90	33.52	32.49	0.83	8.70
Dense Forest	2.92	1.56	1.49	1.82	11.08	0.49	4.73
Hill Shadow	2.62	2.78	2.58	40.28	30.32	0.57	6.63
Mean Variance of total classes	13.33	29.95	58.44	32.91	88.30	0.83	55.92

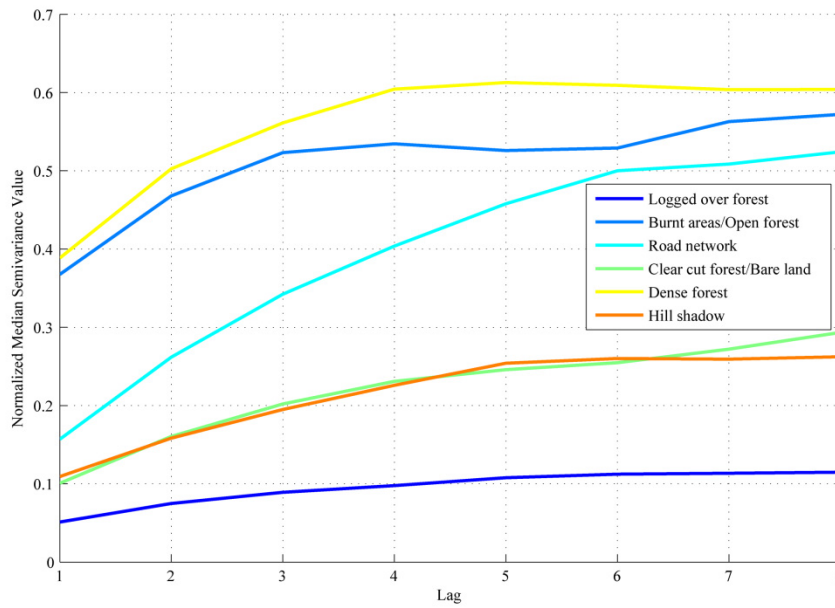
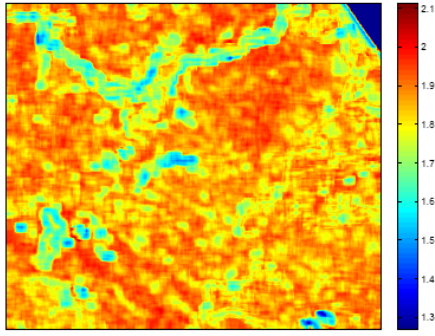


Fig. 3.4 Variogram plots of training data show the spatial variability of land cover classes under study

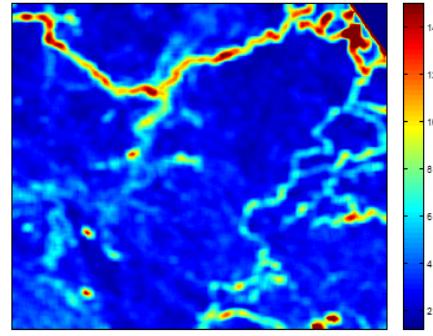
The estimated texture layers show that fractal dimension texture has lower values for non-forest areas and higher values for forested lands. Unlike the fractal dimension texture, other texture features show lower and higher values for forested- and non-forested lands, respectively (Fig. 3.5). The distribution of semivariance values showed that semivariogram, pseudo-cross madogram and pseudo-cross semivariogram showed few differences among forest classes. These texture layers might be less useful to improve the accuracy of spectral data classification. The texture derived using madogram, rodogram and fractal dimension, on the other hand showed better discriminations among forested classes, as well as between forested and non-forested classes. The spectral data classification might gain some benefits by these additional texture features.

3.5.2 Image Classification

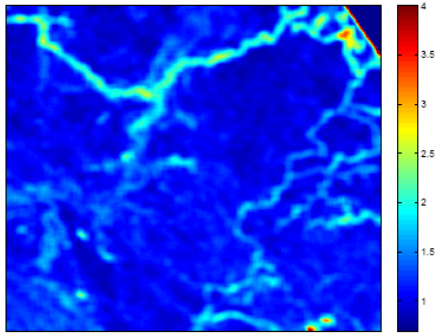
Image classifications were assessed using confusion matrices and we found the average accuracy of MLC was 79.28% which was higher than the MLP-NN and SVM methods with 76.77% and 77.01% of accuracies, respectively. Introducing one texture layer in the spectral data classification, we found fractal dimension has increased the classification accuracy of 2% for SVM classification and up to 4% for MLP-NN classification. In the case of maximum likelihood, fractal dimension contributed no significant improvement in the classification accuracy (Fig. 3.6).



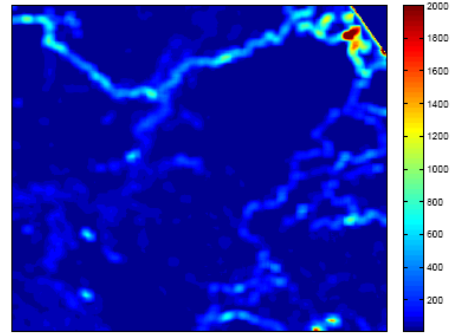
(a) Fractal Dimension



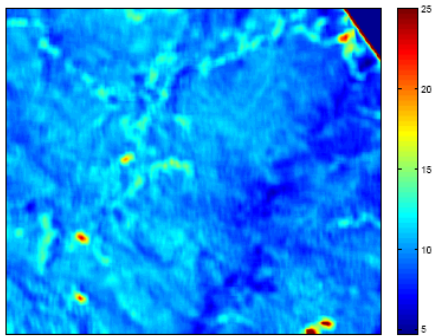
(b) Madogram



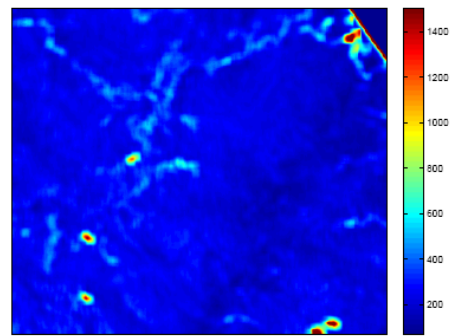
(c) Rodogram



(d) Semivariogram

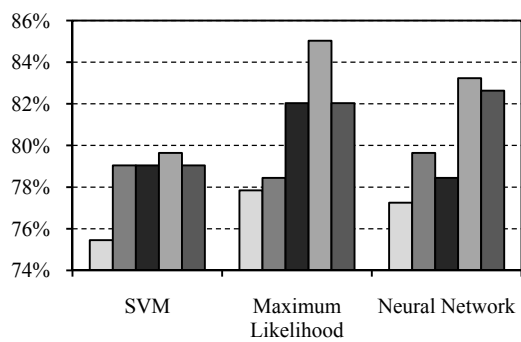


(e) Pseudo-Cross Madogram

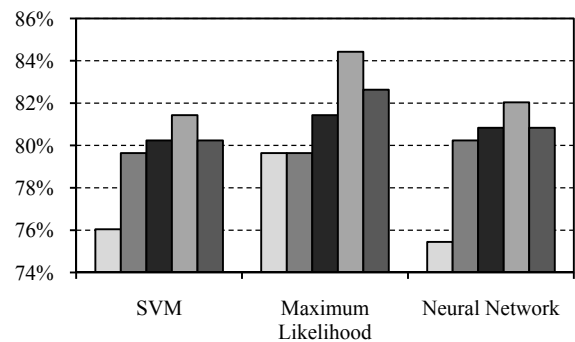


(f) Pseudo-Cross Semivariogram

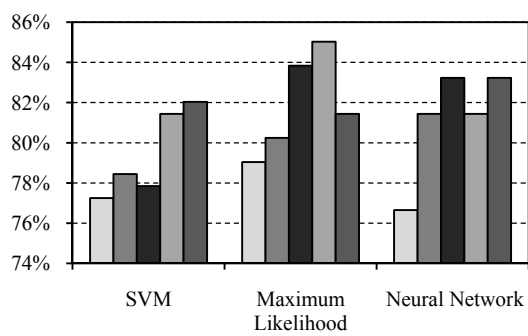
Fig. 3.5 Texture features derived from Landsat ETM data using spatial variability measures of Geostatistics and fractal dimension approach



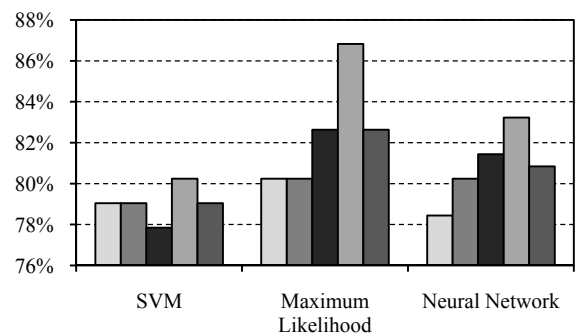
(a) Landsat ETM Band 345 and texture feature



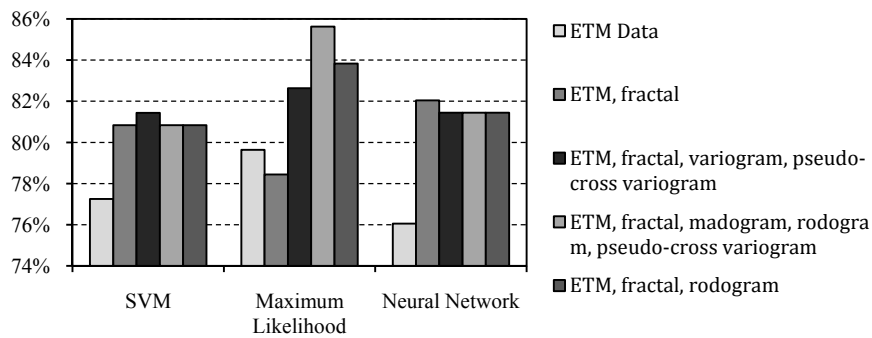
(b) Landsat ETM Band 1345 and texture feature



(c) Landsat ETM Band 1247 and texture feature



(d) Landsat ETM Band 12347 and texture feature



(e) Landsat ETM Band 123457 and texture feature

Fig. 3.6 Accuracy assessments of classification results using different texture data. Each bar represents different variations of texture features used in the classification as explained in the figure legend

Experiments initiating two variations of texture features were performed, we found fractal dimension and rodogram provided better accuracy as compared to other texture variations. Combined with both textures, classification accuracies of MLC, MLP-NN and SVM reached 82.51%, 81.80% and 80.24%, respectively. The mean accuracy of these methods was 81.52% higher than the mean accuracy of spectral data classification (77.69%).

The experiments featuring more texture layers were conducted, and we found that by introducing fractal dimension, variogram and pseudo-cross variogram the accuracies of MLC, MLP-NN and SVM increased up to 82.51%, 81.08% and 79.28% respectively with an average of 80.96% accuracy. Four texture features, namely fractal dimension, madogram, rodogram and pseudo-cross variogram, and Landsat data were employed and we obtained the classification accuracies of MLC, MLP-NN and SVM were 85.39%, 82.28% and 80.72% with an average of 82.79% accuracy. In general, the experiments conducted in this study revealed that addition of geostatistics texture features certainly can improve classification accuracy of spectral data. Performance of geostatistics texture features have been tested on a statistical classifier, i.e. MLC, and two non-statistical classifiers, i.e. MLP-NN and SVM. Application of geostatistics texture features on these classifiers demonstrated a positive result on the classification accuracy.

3.6 Discussion

3.6.1 Estimation of texture features moving window size

Geostatistics texture data performed well in the classification when the data were used as extra information and were combined with spectral information. The attempts applying only geostatistics texture as the primary input for the classification yielded a very poor accuracy. This is because the texture data, considering neighboring values of a single pixel, reduced coarseness of unsimilar and small substance pixels in a neighborhood, which in turn reduced the classification accuracy.

Selection of a moving window size is probably the main issue that is important for extracting useful information of texture features from satellite data. Similarities of neighboring pixels which mainly represent certain land cover types only provide useful texture information if such texture is estimated at proper scale. This means before estimating texture data, a proper moving window size representing an object (i.e. certain land cover type) on the ground should be determined. Using geostatistics, spatial variability of adjacent pixels was explained by semivariance parameters, namely sill and optimal lag distance (range), and the size of moving window was selected accordingly. To observe spatial correlation of each land cover type, semivariance values of the land cover types were

calculated at a 7×7 window size using training dataset up to 8 lag distance (or pixels), and the result is shown in Fig. 3.4. This figure shows that neighboring pixels of each land cover type reached maximum sill within 5 lags (or range is equal to 5 pixels), except for road network class which reached maximum sill of more than 5 lags. For all land cover types except road network, spatial correlation appeared within 5 pixels distance. However, the fact that road network has range larger than other land cover types, and texture data of this respective class should be estimated separately is beyond the scope of this study. Therefore as a compromise with other land cover types, geostatistics texture was estimated using 5×5 window size. This finding is confirmed by other studies that found useful spatial information from geostatistics texture features in a relatively small moving window (Chica-Olmo and Abarca-Hernandez, 2000; Miranda and Carr, 1994).

3.6.2 Behavior of texture features

The behavior of geostatistics texture was studied introducing variations of texture layer in the spectral data classification. From the accuracy assessment of confusion matrices, we obtained that assigning single texture data, semivariogram and pseudo-cross semivariogram performed less satisfactorily than fractal dimension, rodogram and madogram. This is due to the nature of semivariogram and pseudo-cross semivariogram, estimating the mean sum of square of semivariance for all observed lag distance, is less friendly with the presence of data outliers. Compared to semivariogram and pseudo-cross semivariogram, semivariance in rodogram were estimated by computing the mean sum of square root of absolute pixel values difference, while in madogram semivariance were estimated by calculating the mean sum of absolute pixel values difference. The two latter approaches have a 'softer' effect to data outliers as well as extreme values as compared to semivariogram and pseudo-cross semivariogram. In fractal dimension, slope of log-log semivariogram fractal dimension was estimated and this texture measure provides higher variance of land cover classes, resulting better separability of training data in the classification.

We also obtained that application of fractal dimension and rodogram in spectral classification provided better accuracy compared to other two-texture variations. Classification using spectral data, rodogram and fractal dimension yielded considerably good results, with an average accuracy of 81.52% higher than the best three-texture variation, i.e. using fractal dimension, semivariogram and pseudo-cross semivariogram, with 80.96% average accuracy (Tab. 3.2). To visualize how rodogram and fractal dimension can improve data separability, training data of these texture measures were plotted on two dimensional feature spaces as described on Fig. 3.7. Spatial distribution of training data plotted on rodogram and fractal

Tab. 3.2 Overall accuracy assessment (OAA) and Kappa statistics of spectral and geostatistics texture data classification

Input Data	SVM†		Maximum Likelihood†		MLP Neural Network†		Accuracy Average‡
	OAA (%)	Kappa	OAA (%)	Kappa	OAA (%)	Kappa	
ETM*	77.01%	0.7242	79.28%	0.7517	76.77%	0.7210	77.69%
ETM*, fractal	79.40%	0.7528	79.40%	0.7531	80.72%	0.7676	79.84%
ETM*, fractal, variogram, pseudo-cross variogram	79.28%	0.7513	82.51%	0.7904	81.08%	0.7724	80.96%
ETM*, fractal, madogram, rodogram, pseudo-cross variogram	80.72%	0.7684	85.39%	0.8247	82.28%	0.7866	82.79%
ETM*, fractal, rodogram	80.24%	0.7627	82.51%	0.7904	81.80%	0.7809	81.52%

*ETM here refers to different Landsat band variations, i.e. band 345, band 1345, band 1247, band 12347 and band 123457 (Multispectral bands)

†The overall accuracy assessment (OAA) and Kappa statistics are the average accuracy of respective values resulted either from the SVM, Maximum Likelihood and MLP Neural Network methods using variation of ETM input data mentioned in (*)

‡The accuracy average is calculated from the OAA of the SVM, maximum likelihood and MLP Neural Network

dimension data clearly shows how these texture measures could ease separation of training data for each land cover type, except logged over forest. Nevertheless, this is a good indicator showing that rodogram and fractal dimension can improve the classification results. Spectral classifications applying fractal dimension and rodogram, as expected, provided a high mean accuracy (Tab. 3.2).

The accuracy assessment of error matrices revealed that dense forest, burnt area/open forest and road network classes were less misclassified when fractal dimension and rodogram applied in spectral classification. This was clearly exhibited by the feature space of rodogram and fractal dimension texture where training data of those classes were discriminated by both textures (Fig. 3.7).

Finally, we tested four-texture features in spectral classification and the accuracy assessments showed that fractal dimension, madogram, rodogram and pseudo-cross variogram outperformed other texture variations with the accuracy average of 82.79%. These input data significantly improved the overall classification accuracy and the accuracies of each land cover were more than 50%. In general, these texture features improved the accuracies of less accurate land cover classes (i.e. classes with the accuracy of less than 60%) up to 20% in average.

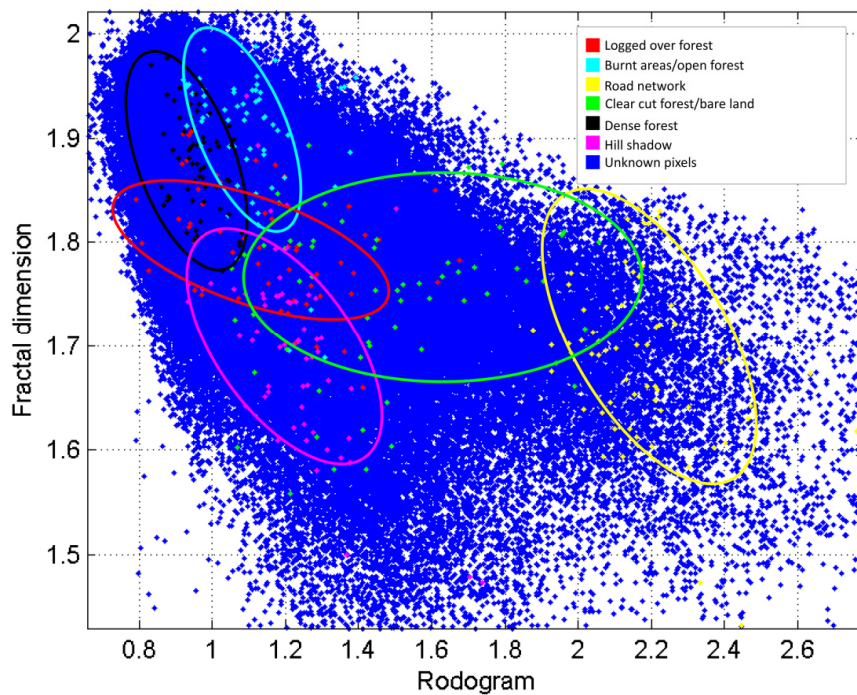


Fig. 3.7 Feature Space of rodogram and fractal dimension showing spatial distribution of total pixels on the study area and training data of each land cover class

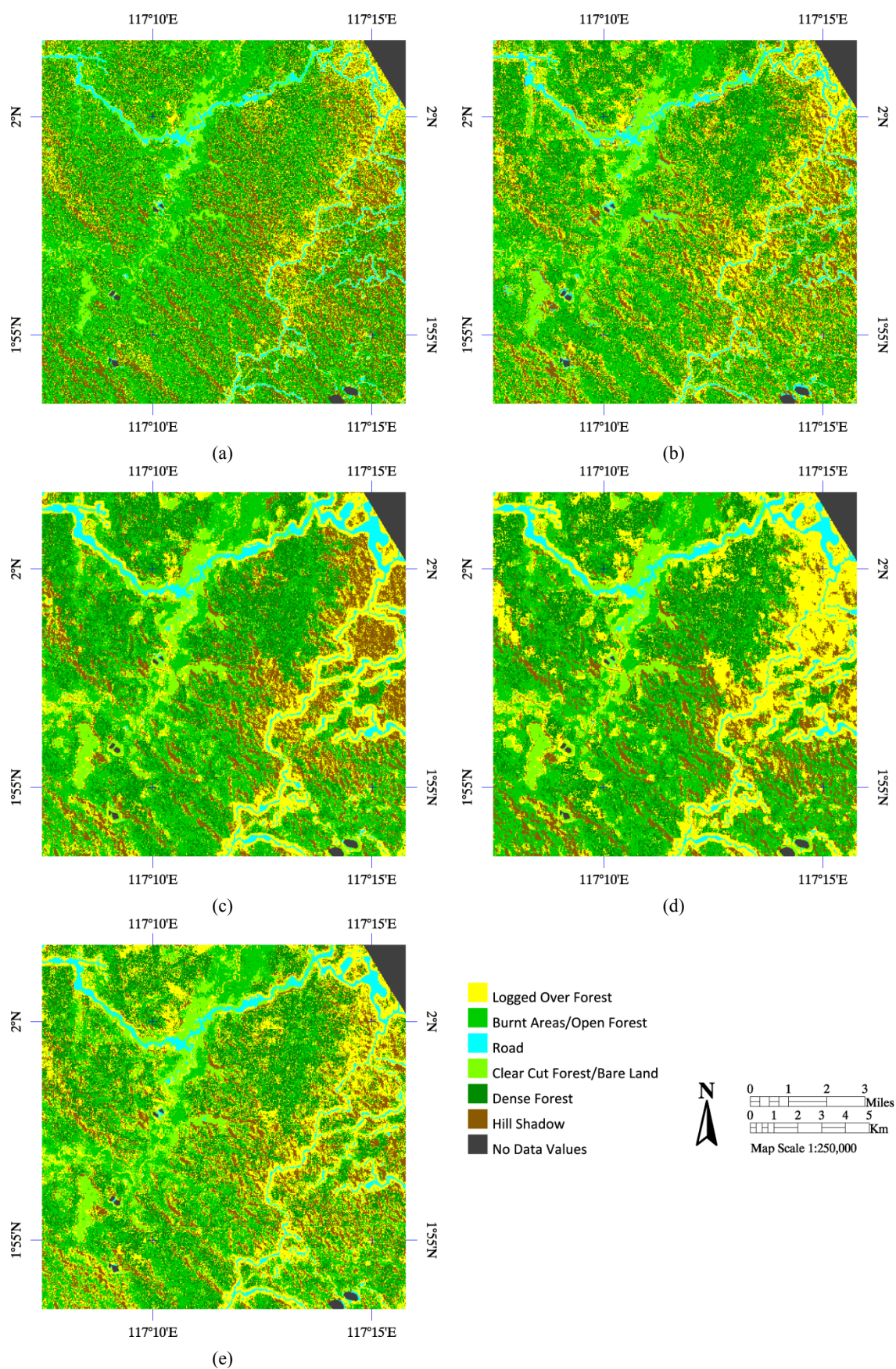


Fig. 3.8 Subsets of maximum likelihood classification using the input data of (a) ETM 12347, (b) ETM 12347 and fractal dimension, (c) ETM 12347, fractal, variogram, pseudo-cross variogram, (d) ETM 12347, fractal, madogram, rodogram, pseudo-cross variogram, and (e) ETM 12347, fractal dimension, rodogram

Texture features, considering spatial arrangement of neighboring pixels, have the potential to improve the classification of linear objects, such as road network. This study found that madogram and fractal dimension were useful for extraction of road network class, provided the evidence that this particular object was clearly exhibited in both texture layers (Fig. 3.5).

Based upon visual observation on the classification results (Fig. 3.8), geostatistics texture data has reduced 'salt and pepper' effects, in which some pixels of certain land cover types are scattered and scratched among other land covers. These effects, which make classified image becomes very heterogeneous, are very common on the classification results of spectral data and cause delineation of land cover boundaries becomes more difficult. The salt and pepper effects were reduced especially when four texture layers, namely fractal dimension, madogram, rodogram and pseudo-cross variogram were assigned as extra information in the spectral classification. Apparently considering the value of neighboring pixels, geostatistics texture lowered the gray level differences of adjacent pixels; thus reduced the salt and pepper effects and improved the classification accuracy.

3.7 Conclusions

This study found that texture layers estimated using geostatistics have significantly improved spectral data classification of Landsat 7 ETM+ image. Texture layers were computed using 5×5 of moving window since the properties of land cover types in the study area were better described using this window size. Selection of the window size was very important as useful texture information could be provided if the moving window had similar size as the actual scale of the features. Image classifications were conducted using MLC, SVM and MLP-NN methods and the results showed that the performance of geostatistics texture features was very effective in improving the classification accuracy and the characteristics of classified image.

We found that texture features derived using rodogram, madogram and fractal dimension have more information than other texture features. The highest accuracy is achieved when fractal dimension, madogram, rodogram and pseudo-cross variogram were applied as additional information of the classification of spectral data.

3.8 Acknowledgments

The Landsat 7 ETM+ image and ground truth data were collected when the first author joined to the MONCER Project during his master study at the International Institute for Geo-Information Science and Earth Observation the Netherlands. Therefore, the first author

would like to thank Dr. Ali Sharifi and Dr. Yousif Ali Hussein, who helped in data collection used for the purpose of this study.

Chapter 4

Improved Strategy for the Estimation of Above Ground Biomass using Non-Destructive Approach[†]

ABSTRACT

We propose a non-destructive approach using a neural network method on optical remote sensing data to estimate stand volume and above ground biomass (AGB). The Levenberg-Marquardt algorithm was selected for training the network, as it was found faster than standard back-propagation algorithm. Surface reflectance of Landsat ETM+, vegetation indices, and SRTM DEM were used to predict stand volume per hectare combined with number of stems data. The combination of satellite and field observation data explained the predicted variable better than did the individual data. AGB was converted from the estimated stem-volume using the Brown – Lugo and Fearnside models. The results for the Brown – Lugo model were similar to those found in previous studies. The estimated AGB was extrapolated over the study area using Kriging. We related the spatial distribution of AGB with aggregated land cover map of the study area and found higher AGB in mature forests than in regenerating- and open/riverine forests.

Keywords: above ground biomass, stand volume, remote sensing, neural networks, Levenberg-Marquardt, Kriging interpolation

[†] This chapter is based on:

Wijaya, A., R. Gloaguen and H. Heilmeyer. *Non-Destructive Approach Development using Levenberg-Marquardt Neural Network for Estimating Stand Volume and Above Ground Biomass in Secondary Tropical Forests of Indonesia*. International Journal of Remote Sensing (In review).

4.1 Introduction

In environmental study, biomass assessment is important for many purposes. It is aimed at two major objectives: (1) for resource use and (2) for environmental management (Parresol, 1999). It is important to determine how much fuel wood or timber is available for use. Thus, one needs to know how much biomass is available at one given time. In environmental management, biomass quantification is important to assess the productivity and sustainability of forests. Biomass and its temporal change is also an important indicator in carbon sequestration. For this purpose, one needs to know how much biomass is lost or accumulated over time. Consequently, the amount of carbon sequestered can be inferred from the biomass change since 50% of the forest dry biomass is carbon (Losi et al., 2003). The Kyoto protocol requires transparent reporting of forest removal and accumulation (biomass change). This implies the requirement of precise procedures to quantify forest biomass. Forest biomass in general may be divided into above ground- and below ground biomass. The above ground biomass (AGB) consists of all living biomass above the soil including stem, stump, branches, bark, foliage and seeds or fruits. Below ground biomass consists of all living roots. Fine roots of less than (suggested) 2 mm diameter are excluded because these often cannot be distinguished empirically from soil organic matter or litter (FAO, 2004). While the BGB dynamically changes once a forest area is disturbed, the AGB has more contributions to total forest biomass, which are often around 75% - 85% (Houghton et al., 2001). This study focuses to estimate the AGB using an improved strategy combining remote sensing and in situ data. Within this study the AGB is defined as the total amount of above ground living organic matter of forest stands expressed in oven-dry weight per unit area.

There are basically three approaches for AGB mapping, which are based on field measurements, Geographic Information Systems (GIS), and remote sensing (RS) data (Lu, 2006). The first approach is accurate but very costly and time consuming, as a destructive sampling is required (De Gier, 2003). Recently, a remote sensing based approach has become an efficient and popular technique for biomass assessment, especially when in situ data are unavailable. Previous studies explored the potential of optical satellites (Foody and Cox, 1994; Houghton et al., 2001; Lu et al., 2004), laser scanner (LIDAR) data (Lefsky et al., 2002), or radar data (Austin et al., 2003; Rauste, 2005) to estimate the AGB. The state of the arts of remote sensing based biomass estimate was based on LIDAR data, which could measure tree height and volume of canopy cover accurately. This data is suitable for mapping of the local scale of forest biomass, which needs high accurate estimation. However, temporal resolution of this data is low as data acquisition is expensive, and if available, the LIDAR data usually

covers only very small portions of a forest region. Moreover, this data are mostly available for boreal and temperate forests, which is not a concern of this current work. For regional and global biomass assessments, remote sensing studies usually used moderate resolution of optical remote sensing data, e.g. Landsat ETM and MODIS sensors, or Synthetic Aperture Radar (SAR) data which are now widely available, including those in the tropical rainforests (Hajnsek et al., 2009; Kuplich et al., 2005; Luckman et al., 1996).

For over a decade the biomass study in tropical environments mainly addressed the challenge and problem in the Amazonian forests (Brown and Gaston, 1995; Kuplich et al., 2005; Lu et al., 2004; Sales et al., 2007; Steininger, 2000). The AGB model for this particular region is already well established (Houghton et al., 2001), whereas in another tropical forest region, e.g. South East Asian rainforests, only few studies, particularly those applying remote sensing approach, were documented. (Foody et al., 2003; Hajnsek et al., 2009). Indonesia possesses the third largest tropical rainforests around the globe after the Brazilian Amazon and tropical forest of Republic of Congo (FAO, 2003). Indonesian rainforests, however, were greatly devastated, especially during El Nino and severe droughts and massive forest fires in the 1990s (Stibig and Malingreau, 2003). Persistence forest encroachment mainly triggered by timber harvesting and agricultural land conversion, complemented with the force of population increase were the major problem for the forest region (Sist and Nguyen-Thé, 2002). Once the forests were deforested, their economic values become lower, and the area is prone to be converted as agricultural lands or other uses. In relation to total biomass abundance, the carbon storage capacity in these deforested lands was heavily depleted. Gradient of forest successional stages markedly appeared in the degraded forests, and were generally expressed as mature forests, regenerating forests, and open forests. The AGB assessment in the different forest gradient is required, so that the biomass density from different forest physiognomy can be estimated.

More specifically, this study is conducted in a lowland *dipterocarp* forest of the Kalimantan Island. For the last two decades, most of tropical lowland forests in this island were either heavily logged or severely degraded and deforested. Evaluation of biomass estimate that can easily be applied without using destructive sampling is required, because this approach is useful to monitor the biomass dynamics in regular basis. A study reported in this forest region estimated the AGB by destructively sampling of more than 40 trees (Samalca, 2007), but failed to find correlations between the biomass and satellite data, which if successfully implemented, this approach may have potential to complement the conventional approach and might be used to assess the biomass regularly. Most studies, even if they found a correlation between the AGB and RS data, usually used a regression method (Lu et al., 2004;

Sales et al., 2007), which is questionable, especially if the input data were from different sources (Wijaya and Gloaguen, 2008). This is because the regression method is a parametric approach that requires normal distribution of the input data to perform properly, and the utility of multi source data may certainly violate this normality assumption. Unlike previous study, we proposed a nonlinear least square function based on a neural network method, which was trained using the Levenberg-Marquardt algorithm to estimate the forest biomass (Hagan and Menhaj, 1994). This algorithm was more efficient than the normal back-propagation method in training the neural network (Wijaya and Gloaguen, 2008; Wijaya et al., 2008a). Calibrated spectral data of Landsat 7 ETM+ image, vegetation indices, DEM and number of stems data were used as predictors in estimating forest stand volume, and the AGB was predicted from the stand volume estimate applying two biomass conversion models (Brown and Lugo, 1992; Fearnside, 1997). We expect this improved strategy for predicting the forest stand volume may result in more favorable AGB estimation. The estimate results are compared with previous studies and the results are discussed. Additionally, the spatial distribution of the AGB is calculated using Kriging interpolation method, and the effects of land cover change and current policies on the biomass density are explained.

4.2 Data and Methods

4.2.1 Field data analysis

We used a temporary sample plot (TSP) field inventory data acquired in 1997 – 1998 (Tab. 4.1). The forest concession management implemented a stratified systematic approach for collecting the TSPs, which were diagonally set up along six transects following a SE to NW direction. Each transect was ± 30 km long; the distance between transects spanned between 3.4 - 5 km. The TSPs were bordered by the Kelai River in the east and the Siduung River in the west. The spacing between each TSP was 100 m on average. In each TSP, tree measurements (i.e., diameter at breast height (dbh) and number of stems) were conducted using three nested subplots, i.e., 0.125 ha for dbh >50 cm, 0.04 ha for dbh 20-49 cm, and 0.0125 for dbh 10-19 cm. Coordinates of sample plot center were also recorded using GPS. The total amount of sample was 1512 plots covered approximately 0.24% of the total area of 81,224 ha. From the tree measurements, the basal area of each tree was calculated and stem volume per hectare was estimated using stand volume table.

Tab. 4.1 Descriptive statistics of transect data

	Number of stems (ha ⁻¹)	dbh (cm)	BA (m ² .ha ⁻¹)	Stand Volume (m ³ .ha ⁻¹)
Mean	264	33.7	14.0	159.3
Min	5	10.0	0.4	1.0
Max	1055	130.0	53.6	628.6
SD	146	10.6	7.4	92.7
N	1512			

4.2.2 Satellite image preprocessing and vegetation indices generation

We used a 30 m resolution Landsat 7 ETM+ image acquired on August 26, 2000 for this study. The image was atmospherically and topographically corrected, resulting in surface reflectance values. There are several ways to conduct atmospheric corrections on remote sensing data (Chavez Jr., 1996; Richter, 1996; Schroeder et al., 2006; Song et al., 2001). Due to unavailability of physical atmospheric data, this study implemented the atmospheric correction based on the dark pixel or dark object subtraction (DOS) method (Chavez Jr., 1996). This approach assumes that in each image scene, dark objects exist; thus, the radiometric correction can be done by subtracting the pixel values of dark objects from all pixels in the image scene. Using DOS technique, only the atmospheric effects caused by atmospheric scattering were corrected, as this method does not model a complex atmospheric absorption effect in its equation.

Topographic corrections were conducted to reduce the compensation of solar illuminations due to the irregular shape of terrain. We used a Digital Elevation Model (DEM), which was geo-referenced into the same coordinate system as the satellite image so that the slope angle and aspect of each pixel position could be derived. The topographic corrections were carried out using C-Correction method assuming Lambertian surface on the earth surface (Riaño et al., 2003).

From the corrected image, we calculated vegetation indices, i.e., Normalized Difference Vegetation Index (NDVI), Simple Ratio (SR), Enhanced Vegetation Index (EVI), and Atmospherically Resistant Vegetation Index (ARVI). Details about vegetation indices can be found in remote sensing references (Jensen, 1996; Mather, 2004; Richards, 1993). We also used DEM from Shuttle Radar Topography Mission (SRTM) data to derive elevation, slope, and aspect maps over the study area.

4.2.3 Estimation of stand volume

Prediction of stand volume was carried out using a neural network method combining remote sensing and field observation data. This combination could violate normality assumption and reduce the predictive ability of parametric methods, such as regression (Berry and Feldman, 1985). With a neural network, the data do not need be normally distributed, as this method has characteristics between parametric and non-parametric models (Cohen and Cohen, 1983). To predict the stand volume, we experimented using three input variations, namely (1) number of stems data (i.e., field data), (2) spectral values, vegetation indices, and DEM (hereinafter satellite data), and (3) combination of both data.

Neural networks work by aggregating the weighing of the input values and applying a threshold function into those values. This training process is sometimes very time-consuming, and the selection of a faster and more efficient training algorithm is preferable. For remote sensing applications, back-propagation is probably the most widely used training algorithm for the neural networks (Atkinson and Tatnall, 1997). Unfortunately, this algorithm is very time consuming, and to yield an accurate result the training parameters should be selected carefully.

Within this study the network was trained using the Levenberg-Marquardt algorithm, which we found more efficient than back-propagation algorithm. The Levenberg-Marquardt is basically an iterative technique that locates the minimum of a function expressed as the sum of squares of nonlinear functions (Hagan and Menhaj, 1994). It is a standard technique for nonlinear least-squares problems and can be thought of as a combination of steepest descent and the Gauss-Newton method. When the current solution is far from the correct one, the algorithm behaves like a steepest descent method: slow, but guaranteed to converge. When the current solution is close to the correct solution, it becomes a Gauss-Newton method. Given a function $F(x)$ that is a sum of squares of nonlinear functions,

$$F(x) = \frac{1}{2} \sum_{i=1}^m [f_i(x)]^2 \quad \text{Eq. 4.1}$$

to find the minimum of function $F(x)$, the Jacobian of $f_i(x)$ is denoted $J_i(x)$, and the Levenberg-Marquardt method searches in the direction given by the solution p to the equation

$$(J_k^T J_k + \lambda_k I) p_k = -J_k^T f_k \quad \text{Eq. 4.2}$$

where λ_k are nonnegative scalars and I is the identity matrix. For some scalar Δ related to λ_k , the vector p_k is the solution of the constrained sub-problem of minimizing $\|J_k^T p + f_k\|_2^2/2$ subject to $\|p\|_2 \leq \Delta$.

Before training the network, we divided the input data into three subsets, i.e., 60% of the data were used as a training dataset, and 40% of the sample data were divided into two subsets and used as validation and test data. Training datasets were used to train neural networks, while validation datasets were used as a control of the training process. When the accuracy of validation data increased, the training was stopped to avoid overtraining of the network. A test dataset was used as independent data to assess the prediction accuracy. We tested different learning parameters, such as number of hidden neurons and number of iterations, and averaged the estimated value. We compared the results with back-propagation neural network and linear regression and assessed the accuracy of each method in predicting the stand volume.

4.2.4 Above ground biomass (AGB) prediction

We estimate the AGB using the stand volume – biomass equation proposed by Brown and Lugo (1992) and Fearnside (1997). Both equations were developed for tropical moist forests of the Brazilian Amazon, and were commonly used for the biomass assessment over this forest region (Houghton et al., 2001; Sales et al., 2007). The Brown – Lugo AGB (AGB_{BL}) is explained as follows:

$$AGB_{(BL)} = SB \cdot BEF \cdot (1 + (0.09 + 0.21)) \quad \text{Eq. 4.3}$$

where SB is the stemwood biomass, and is defined by

$$SB = Volume \times VEF \times WD \quad \text{Eq. 4.4}$$

with $Volume$ ($m^3 \cdot ha^{-1}$) obtained from forest inventory data. VEF (volume expansion factor to account for trees smaller than the minimum diameter measured) was 1.25 for dense forests and 1.5 for other than dense forests. WD is the wood density ($0.69 \text{ ton} \cdot m^{-3}$ as a weighted average of wood density), and BEF (biomass expansion factor to account for biomass in addition to stemwood biomass) varied as a function of stemwood biomass (SB) as follows (Brown and Lugo, 1992).

$$\text{for } SB < 190 \text{ Ton ha}^{-1}, BEF = e^{[3.213 - 0.506 \cdot \ln(SB)]}$$

$$\text{for } SB > 190 \text{ Ton ha}^{-1}, BEF = 1.74 \quad \text{Eq. 4.5}$$

In Eq. 4.3, $SB \times BEF$ is the conversion of volume to above ground biomass originally proposed by Brown and Lugo (1992). The constants 0.09 and 0.21 in the formula were introduced by Houghton et al. (2001) to include below ground biomass and dead above ground biomass.

The second conversion equation, which suggested modification of the first model, was proposed by Fearnside (1997):

$$AGB_{(FS)} = SB \times BEF \times (1 + CF) \quad \text{Eq. 4.6}$$

where $CF = 96.2\%$, which represents the sum of various correction factors (lianas = 5.3%; trees smaller than 10 cm DBH = 12%; tree form factor = 15.6%; trees between 30 and 31.8 cm DBH = 3.6%; hollow trees = -6.6%; bark = 0.9%; palms = 2.4%; below ground biomass = 33.6%; dead above ground biomass soil = 31%; and other components = 0.2%) (Sales et al., 2007). In this study, correction factors from Eq. 4.3 - Eq. 4.6 were assumed constant, i.e., not spatially variable. Comparison with the results of previous studies (Brown, 1997; Samalca, 2007) was conducted and the results were discussed. We used biomass and AGB terms interchangeably in this study.

4.2.5 Spatial distribution of the AGB

AGB estimate was mapped over the study area using the Kriging method, as we assumed the values were spatially correlated and the stationarity assumptions were fulfilled. We transformed the AGB values using log-transform, so that the data were more likely to be normally distributed. An empirical Gaussian variogram model formulated as,

$$\gamma(h) = c \left\{ 1 - \exp\left(-\frac{h^2}{r^2}\right) \right\} \quad \text{Eq. 4.7}$$

where c is the sill, h is the lag number, and r is a sill parameter, was fitted to the spatial variability of the data. The semivariogram modeling showed that no special pattern of spatial variability changed within a certain direction; i.e., no anisotropy appeared. Thus, an omni-directional variogram was selected.

Ordinary Kriging was selected in interpolating the AGB, assuming an unknown constant mean. We assumed that as the predicted location moved farther from the sample data, those points would have less spatial auto-correlation. Thus, interpolation of unknown values considered neighboring pixels within certain distance and direction (i.e., in the NE-SW and NW-SE direction). We selected 40 neighboring pixels for the interpolation of each 'unknown' value. The error estimate was calculated and assessed.

4.3 Results and Discussion

4.3.1 Structure of forest stand properties

In total we analyzed 13050 stands divided into 1512 sample plots. Distribution of measured trees showed that >95% of the trees had a diameter between 10 – 80 cm and the mean dbh of measured trees was 33.7 cm. However, the mean volume of these trees was 5.6 – 27.3 m³.ha⁻¹, which was very small compared to the volume of trees with dbh >80 cm (mean volume 49.3 m³.ha⁻¹) (Tab. 4.2).

The data showed that >70% of sampled trees had dbh <40 cm; however, there was a slight difference between the number of trees with 21< dbh <40 cm and trees with 41< dbh <80 cm. This indicated that many sample plots were dominated by old secondary forests or primary forests. A primary forest has fewer trees but more diverse vegetation than a secondary forest, resulting in greater ecosystem biodiversity. Landsat ETM+ data may be problematic to use with these forests because of the complexity of crown closure structure due to shadow effects (Lu, 2005).

Tab. 4.2 Distribution of forest properties

dbh	Count	Mean dbh (cm)	Mean Basal Area (m ² /ha)	Mean Volume (m ³ /ha)
10-20 cm	5066	14.7	0.8	5.6
21-40 cm	4350	28.9	1.7	18.5
41-80 cm	3035	57.8	2.0	27.3
81-100 cm	435	90.5	3.2	49.3
>100 cm	164	121.9	6.0	96.0
N	13050			

4.3.2 Prediction of stand volume

The predictive ability of Levenberg-Marquardt algorithm ($r=0.641$, $MSE=71.2$ m³.ha⁻¹, $\alpha=0.05$, $n=1512$) was better than that of the back-propagation algorithm ($r=0.608$, $MSE=73.6$ m³.ha⁻¹) and linear regression methods ($r=0.615$, $MSE=73.1$ m³.ha⁻¹) in estimating stand volume (Tab. 4.3). The measurement data showed high variations of stem volume (159.3 ± 92.7 m³.ha⁻¹), which was confirmed by the volume estimates. The high error estimate (~50% of the estimated stem volume) is because the study area is a natural forest dominated with mixed *dipterocarp* sp. vegetations coupled with complex forest structure. The linear regression method ($r = 0.615$) explained the stand volume slightly better than the

back-propagation neural networks ($r = 0.608$), probably due to the problem in selecting proper algorithm parameters.

Tab. 4.3 Prediction results of stand volume

	Mean estimate (m ³ ha ⁻¹)	Min	Max	SD	r	MSE
<u>Stem data</u>						
Linear regression method	159.12	62.61	454.26	54.44	0.586	75.0
Back-propagation neural networks	158.11	60.17	530.76	55.26	0.585	75.1
Levenberg Marquardt neural networks	159.22	36.90	502.64	56.96	0.600	74.1
<u>Satellite data</u>						
Linear regression method	159.04	78.74	338.17	22.58	0.247	89.8
Back-propagation neural networks	158.45	94.01	279.20	24.28	0.247	89.9
Levenberg Marquardt neural networks	154.63	11.87	392.46	31.76	0.247	87.7
<u>Satellite and Stem data</u>						
Linear regression method	159.29	39.01	465.21	57.01	0.615	73.1
Back-propagation neural networks	156.73	40.21	408.22	56.05	0.608	73.6
Levenberg Marquardt neural networks	158.11	9.41	524.19	62.60	0.641	71.2

Training process using Levenberg-Marquardt algorithm was very efficient and requires fewer than 100 iterations to find the optimal solution. The back-propagation method on the other hand, needs at least more than 1000 iterations to obtain similar results. We observed that a 2-layer neural network (1 input, 1 hidden, 1 output layer) was sufficient to solve a general complex problem for this study (we used a maximum of 14 predictors in estimating the stand volume). Higher estimate accuracy was revealed when we used the same- or a similar numbers of hidden neurons to the total input number in the neural networks.

Number of stems per hectare combined with satellite data ($0.608 < r < 0.641$) explain the stand volume better than individual use of data. We also found that the satellite data had a lower correlation coefficient ($r = 0.247$) compared to the stems ($0.585 < r < 0.6$). This was due to the saturation of the satellite data, especially when associated with large stand volume, i.e., mean volume $> 300 \text{ m}^3 \cdot \text{ha}^{-1}$ (Fig. 4.1).

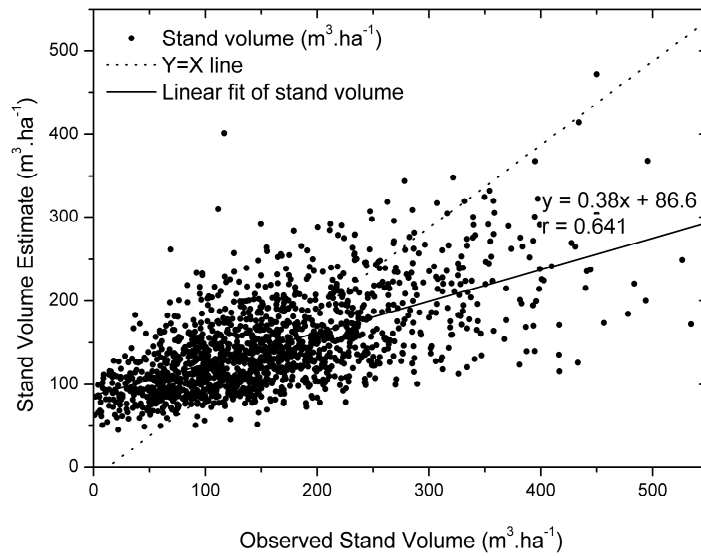


Fig. 4.1 Plot of estimated and measured stand volume

The problem occurred because the interval values of satellite data and vegetation indices were small. The NDVI of green vegetation, for example ranges usually from 0.7 – 1.0. This condition causes the prediction of dependent variable (i.e., stand volume) less accurate when data variance is large (the stand volume varies from 1.0 – 628.6 m³.ha⁻¹).

The coarseness of spatial resolution (30 m for Landsat 7 ETM+) was another issue affecting estimation of the stand volume. One pixel of Landsat ETM+ data covered 900 m² on the ground; with this spatial resolution, variations of the stand volume within a pixel were inaccessible. The satellite data might have been a problem, as our study area was comprised of a highly variable mixture of different forest types and tree species. Spectral energy radiated back from mature or very dense forests and captured by the Landsat ETM+ sensors normally comes from canopy closure or crown cover; tree density is relatively high and often there is no opening in the crown cover. This could be another explanation for the lack of correlation between the Landsat ETM+ data and the predicted stand volume. Nevertheless, this study demonstrated that moderate spatial resolution of optical data (i.e., Landsat 7 ETM+) combined with selected measurement data can improve the estimates of stand volume in highly complex forest environments.

4.3.3 Conversion to above ground biomass

The Brown – Lugo equation predicted lower AGB (mean = 368.3 Mg.ha⁻¹) compared to the Fearnside model (mean = 555.9 Mg.ha⁻¹). This was because of conversion factors difference considered in both models. The Brown – Lugo model included the corrections for dead AGB and BGB in the estimation, while the Fearnside model considered an adjustment factor

related to the physiognomy of various vegetation structures. We found that the Fearnside model overestimated the AGB as confirmed by Samalca (2007) that conducted the study in the Labanan forests estimating the AGB using in situ data and found 335 Mg.ha⁻¹ of dry weight biomass (Fig. 4.2). The Brown – Lugo model obtained more favorable estimate especially if the dead AGB and BGB components considered in the equation were excluded. Using the same dataset, we also predicted the AGB using tree diameter (dbh) as the predictor applying available allometric equation proposed for tropical moist forest environments (Brown, 1997). Using this allometric equation, we predicted in average 155 Mg.ha⁻¹ of AGB, considerably lower than previous estimates. This difference is probably due to above ground biomass in the allometric model was modeled from tree diameter which represents only the biomass of tree stems regardless other parts of forest stands, such as foliage and tree branches, which might also contain biomass. For the purpose of this work, the Brown – Lugo’s AGB estimate was used for the kriging interpolation and compared with land cover map of the study area.

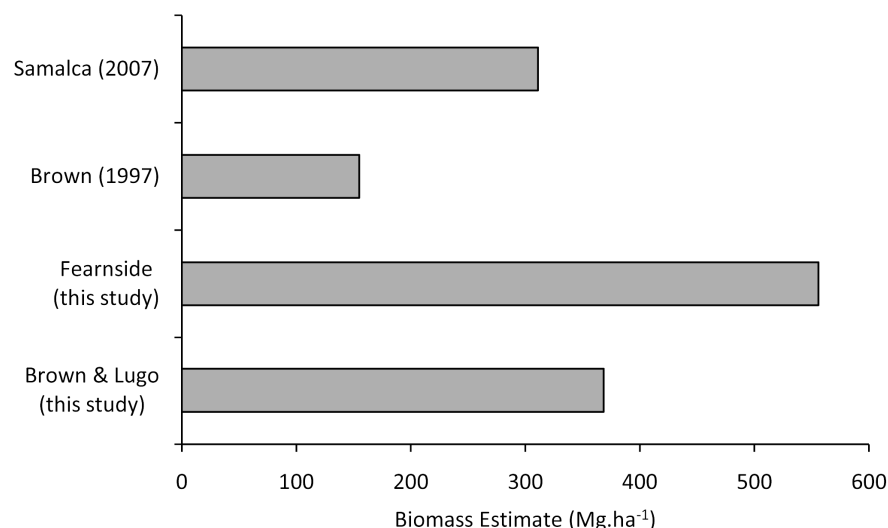
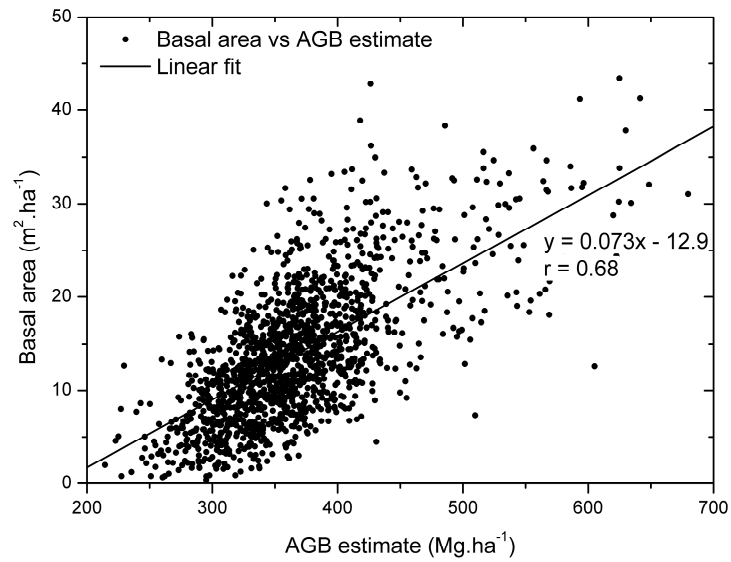
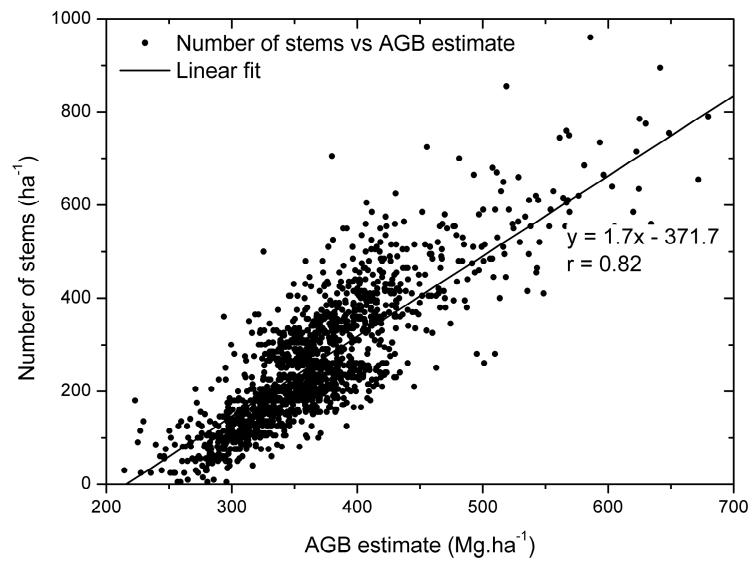


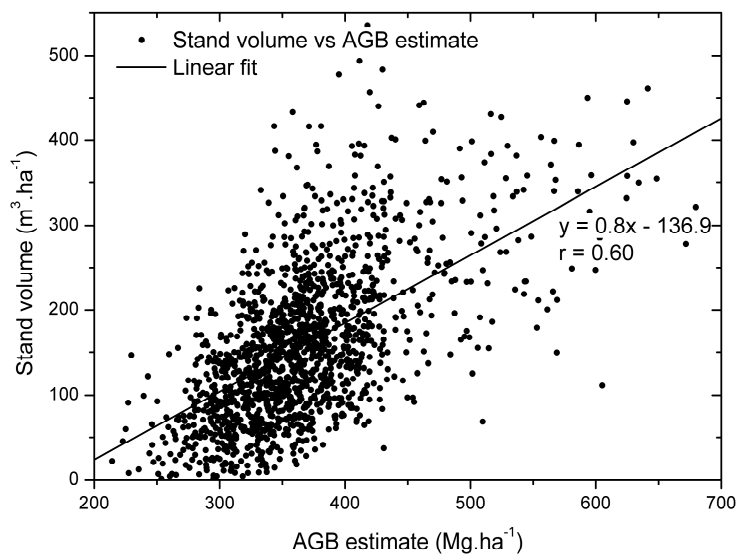
Fig. 4.2 Comparison of biomass estimates. This study used stand volume – AGB equations, which include correction factors for dead AGB, below ground biomass. The study by Samalca and Brown estimated AGB from the allometric equation to relate dbh and AGB



(a)



(b)



(c)

Fig. 4.3 Comparison of biomass estimate with (a) basal area, (b) number of stems per ha, and (c) stand volume

To explain the relationship between biomass and forest stand properties, correlation of those parameters were plotted (Fig. 4.3). Number of stems shows a strong correlation ($r = 0.82$, $\rho < 0.0001$) with the biomass greater than basal area ($r = 0.68$, $\rho < 0.0001$) and observed stand volume ($r = 0.60$, $\rho < 0.0001$). The sample plots were highly scattered, but many of those formed a great cluster of AGB within the range of 250 – 425 Mg.ha⁻¹. This also indicates that most sample data were collected from secondary regenerating forests that expressed moderate biomass density.

The utility of satellite data, e.g. Landsat TM sensor, for estimating the above ground biomass were limited to the saturation of spectral values and vegetation indices, which usually occurs in dense vegetated- or advanced successional forests (biomass density of more than 20 kg.m⁻² or 200 Mg.ha⁻¹), due to the complexity of forest canopy structure and shadow effects (Lu, 2005). Another study by Steininger (2000) even found the asymptotic point of spectral data at lower biomass density (15 kg.m⁻² or 150 Mg.ha⁻¹). Variations of spectral data, vegetation indices and in situ data, i.e. number of stems, are very effective in improving the reliability of the biomass model and to increase the sensitivity of RS data, which in turn may result in more accurate estimation.

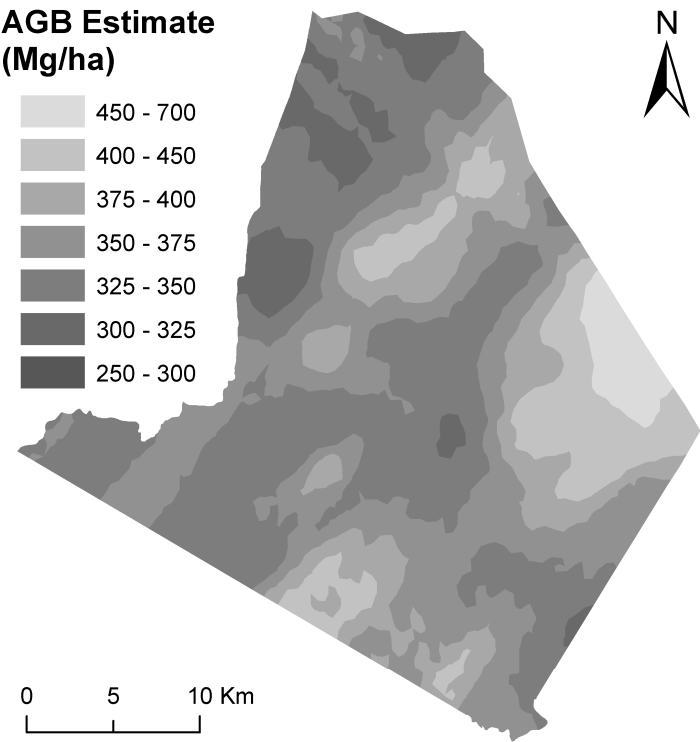
4.3.4 Spatial distribution of AGB

We predicted the biomass estimate over the study area using the ordinary kriging method and found that the AGB density ranged from 214 – 543 Mg.ha⁻¹ with variable error estimate from 55 – 86 ton/ha (Fig. 4.4). Upon kriging interpolation, we found that the error estimate varied from 10 – 43%. These high error estimate were similar with the result of Samalca (2007), who estimated the AGB using a destructive sampling approach and yet found 40% of error estimate.

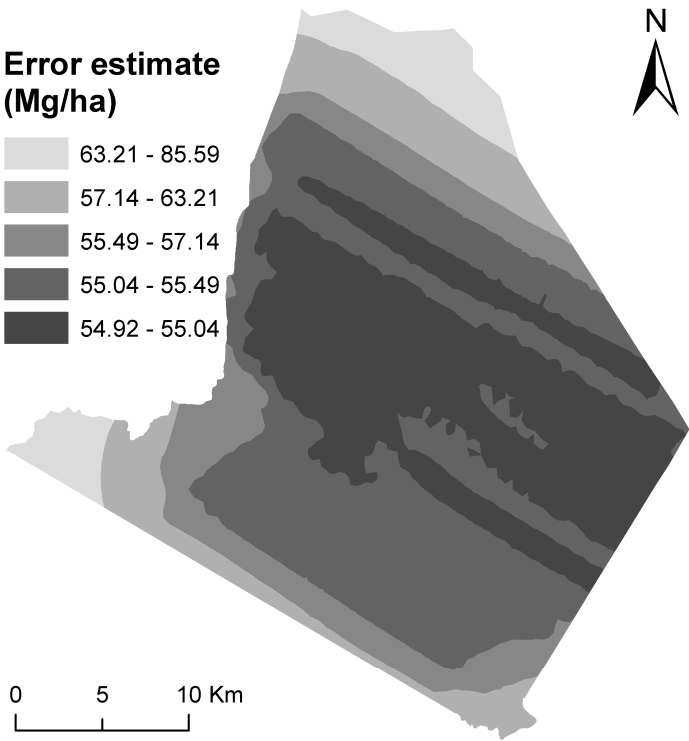
We found that the error estimate increased as the unknown pixels were located more distant than the sampling points (Fig. 4.4b) and confirmed the basis of spatial correlation principle (Journel and Huijbregts, 1981). The error estimate, which is applicable for any spatial based interpolation technique, might be caused by the uneven distribution of sampling intensity among other conditions. In fact, the transect data, which were collected on the ground and originally used for forest inventory, were more concentrated on the central part of the study area.

Another sampling technique, such as stratified random sampling, might be proposed to reduce the error estimate of the kriging method. Unfortunately, it is nearly impossible to collect biomass data only for this study. Despite the error estimate, the interpolated biomass

showed a reasonable relationship with land cover of the study area and is discussed in the following section.



(a)



(b)

Fig. 4.4 Kriging interpolation results (a) and error estimate (b)

4.3.5 Relationship between AGB and land cover classification

Based upon the classified land cover map, we found that the study area had been widely logged selectively and by the forest concession management this area in general was classified into open forests, dense forests, very dense forests, mixed forests and swamp forests (Fig. 4.5).

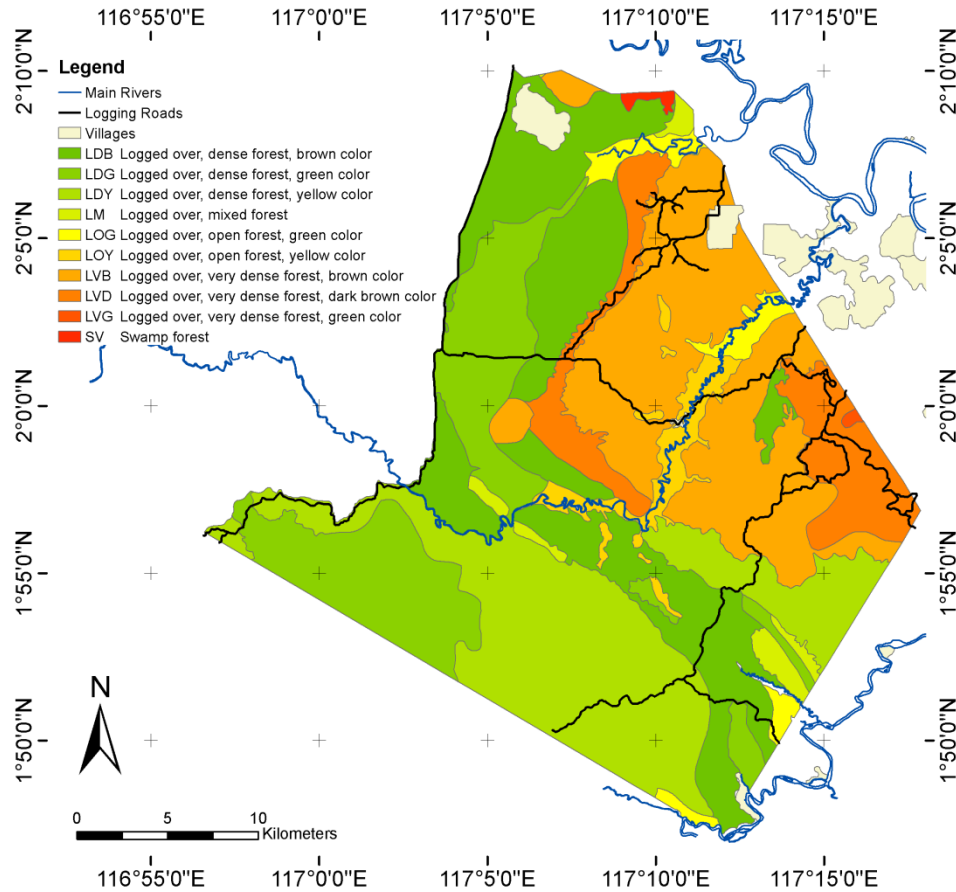


Fig. 4.5 Land cover map of the study area based on the interpretation of Landsat 7 ETM+ images and field survey (Modified from Berau Forest Management Project, 2001)

Considering the gaps of canopy cover, similarity in vegetation structure, and additional information obtained during the field visit, we aggregated some class labels in the land cover map (Fig. 4.5) and merged some similar classes into more general definition, namely open/riparian forests, regenerating forests, and mature forests. The logged over, open forest, green (LOG), logged over, open forest, yellow (LOY), and mixed forest (LM) classes were grouped as open/riverine forests; while logged over, dense forest, brown (LDB), logged over, dense forest, green (LDG), and logged over, dense forest, yellow (LDY) were regenerating forests; and logged over, very dense, forest, brown (LVB), logged over, very dense forest, dark brown (LVD), and logged over, very dense forest, green (LVG) were labeled mature forests. Swamp forest (SV) class was excluded since no sample data were collected in this region.

Open/riverine forests were mainly situated along the rivers or other transportation facility, such as logging roads, giving easy accesses for forest encroachments and expansion of agriculture lands. Regenerating forests were basically secondary forests that had been logged for the last 5 years and were characterized by moderate canopy cover (60% – 75%) and similar tree species distribution to that of an old secondary forest. Mature forests, on the other hand, had a very dense canopy cover which is over 80% and had similar characteristics with undisturbed forests. Some primary forests patches were also found in these areas, which was due to difficult topographic conditions the timber extraction was problematic and the forests remained undisturbed.

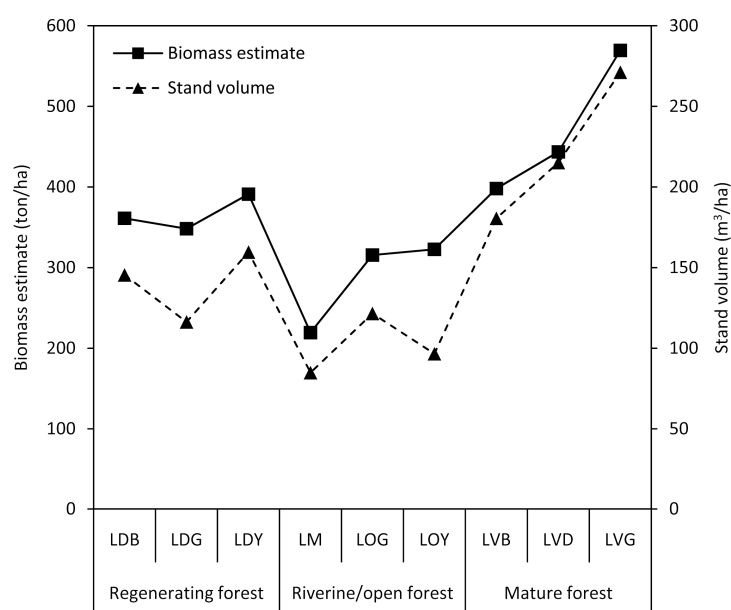


Fig. 4.6 Biomass estimate of each land cover class

The estimation results of stand volume and biomass for each land cover type showed that 175 – 275 m³.ha⁻¹ of stand volume and 400 – 600 Mg.ha⁻¹ of AGB were obtained in mature forests. These values were higher than those obtained in regenerating forests (125 – 160 m³.ha⁻¹ of stand volume, 350 – 400 Mg.ha⁻¹ of AGB) and open forests (75 – 125 m³.ha⁻¹ of stand volume, 200 – 325 Mg.ha⁻¹ of AGB), respectively (Fig. 4.6). The lowest stand volume and AGB were found in a mixed forest (LM), as this forest was dominated by the mixture of swamp vegetation and small trees of the *dipterocarp* species. The appearance of new tree species in logged over forests is much faster than the growing process of tree volume and biomass. The stand volume and AGB of regenerating- and open/riverine forests were noticeable but not significantly different when compared with those in mature forests.

4.4 Policy Implications

Timber harvesting in Labanan forest concession, which had been carried out for over 25 years, stopped in 2003 with the cessation of government licensing. The study area was dominated by secondary and mature forests with the stand volume varies from 9 – 524 m³.ha⁻¹, higher than that in the tropical secondary forests of Amazon, Brazil (Sales et al., 2007). The previous study also reported 225 – 486 Mg.ha⁻¹ of AGB estimated from the secondary forests of the Amazon, slightly lower than the findings of this study (AGB ranges from 214 – 543 Mg.ha⁻¹).

The spatial distribution of AGB and the land cover map were compared. Relatively low forest biomasses and timber stocks were found in forest areas adjacent to the logging roads and main rivers. During the field visit we found many illegal logging sites within 5 km of the roads; the harvested timbers from these areas were easily transported out using trucks, making the areas more susceptible to illegal operations. The rate of illegal logging inside the concession area has increased since implementation of the decentralization policy for the forestry sector in Indonesia, which states that local people can open new agriculture land inside concession boundaries (for up to 100 ha/household). This policy has created serious problems for concession management. Because the opening of new agriculture land supposedly no further than 10 m away from the main roads was ignored, the practice of slash and burn for opening the new agriculture land went further inside undisturbed forests. We obtained lower AGB adjacent to settlement areas as well. Given that two villages shared borders with the forest concession, the excessive slash and burn practice was the major concern for the concession management. The implementation of decentralization policy has triggered a great amount of illegal logging and in general has discouraged forest conservation in Indonesia.

4.5 Conclusion

This study demonstrated that a 30 m resolution of Landsat 7 ETM+ data, combined with selected measurement data, could be used to predict stand volume per hectare. A non destructive approach using remotely sensed data was successfully applied for estimating above ground biomass in a tropical secondary forest. We found 158 ± 63 m³.ha⁻¹ of stand volume and 368.3 Mg.ha⁻¹ of above ground biomass over the study area. The Levenberg-Marquardt neural network outperformed the Back-propagation neural network and Regression methods in estimating stand volume per hectare. Satellite data in combination with number of stems predicted the stand volume better than the individual use of those data. The Levenberg-Marquardt algorithm requires fewer training iterations than the Back-

propagation method to find the optimal solution of the neural network problem. The results of the Kriging method showed that the error estimate of biomass depends on the distribution of sample data. However, this study found a strong relationship between the spatial distribution of AGB and land cover classification. We found higher stand volume and AGB in mature forests than in regenerating- and open/riverine forests.

Chapter 5

Estimation of Stem Volume and Forest Biomass from Remote Sensing and GIS[‡]

ABSTRACT

This study presents the utility of remote sensing (RS), GIS and field observation data to estimate above ground biomass (AGB) and stem volume over tropical forest environment. The AGB density was estimated applying an existing DBH – biomass equation. The estimate was superimposed over the modified GIS map of the study area, and the biomass density of each land cover was calculated. The RS approach was performed using a subset of sample data to develop the AGB and stem volume linear equation models. Pearson correlation statistics test was conducted using Landsat ETM bands reflectance, vegetation indices, image transform layers, Principal Component Analysis (PCA) bands, Tasseled Cap (TC), Grey Level Co-Occurrence Matrix (GLCM) texture features and DEM data as the predictors. To analyze total biomass and stem volume of each land cover, Landsat ETM images from 2000 and 2003 were preprocessed, classified using maximum likelihood method, and filtered with the majority analysis. We found 158 ± 16 m³/ha of stem volume and 168 ± 15 ton/ha of AGB estimated from RS approach. Whereas the field measurement and GIS estimated 157 ± 92 m³/ha and 167 ± 94 ton/ha of stem volume and AGB, respectively. We found a slightly declining trend of total biomass from 2000 to 2003. Remote sensing approach estimated lower biomass abundance than did the GIS and field measurement data.

Keywords: above ground biomass, stem volume, remote sensing, GIS, field observation data

[‡] Wijaya, A., S. Kusnadi, R. Gloaguen and H. Heilmeyer (2010). Improved Strategy for Estimating Stem Volume and Forest Biomass using Moderate Resolution Remote Sensing Data and GIS. Journal of Forestry Research, Vol. 21 No. 1, pp. 1 – 12.

5.1 Introduction

Current information on above ground biomass (AGB) is important to estimate carbon accumulation over a forest region and it is required to study the impacts of forest disturbance on total biomass. The AGB can be estimated using different data and approaches, namely using field observation data (Brown et al., 1989; Brown and Lugo, 1984; Brown and Lugo, 1992), remote sensing (RS) data (Barbosa et al., 1999; Foody, 2003; Roy and Ravan, 1996; Steininger, 2000; Thenkabail et al., 2004), and GIS (Brown and Gaston, 1995; Brown et al., 1994). Field observation approach is known to be the best and the most accurate method, but it is costly and time-consuming as destructive sampling data is required (De Gier, 2003; Lu, 2006). RS and GIS approaches recently become more popular as huge areas can be covered with less efforts and time, with regard to different sensor characteristics and limitations (Houghton et al., 2001; Lu, 2005; Lu, 2006). Estimation of AGB is still a challenging task since the utility of RS and GIS for the biomass modeling is site specific and is highly uncertain (Foody et al., 2003; Houghton et al., 2001). Performance of RS data and combination of field data – GIS in estimating the AGB is presented in this work.

The application of remote sensing data and techniques for AGB prediction have been widely studied, employing optical sensor (Lu, 2005), SAR data (Hajnsek et al., 2005) or LIDAR data (Lefsky et al., 2002). These studies found that state of the art LIDAR data could provide the most accurate result as it allows a deep penetration through forest canopy (Lu, 2006). The utility of polarimetric interferometry SAR data (PolinSAR) for biomass estimation is also widely studied (Hajnsek et al., 2005). This data provides useful information on digital surface model, which can easily be converted into biomass using some inversion models (Cloude et al., 2008; Hajnsek et al., 2005; Isola and Cloude, 2001). Unfortunately, the potential of these data cannot be demonstrated here due to data unavailability. Alternatively, moderate resolution of Landsat ETM data coupled with vegetation indices, image transform layers, PCA, Tasseled caps, Grey Level Co-occurrence Matrix (GLCM) texture features and SRTM DEM were considered.

Different forest disturbance and harvesting regimes could have occurred over a forest area. Once these disturbances were over, forest regenerating processes are started. The intensity of these processes is different for each forest region depending on climate, terrain conditions, soil fertility and nutrient contents, characteristics of pioneer vegetation species, etc. In natural secondary forests, a mixture of different forest physiognomy, e.g. young forest, regenerating forest, old secondary forest, etc., is easily noticed. This study has objectives to estimate AGB and stem volume over different forest succession stages. Because recent studies arrived at different conclusions on the biomass assessment when RS data

were applied applied (Foody et al., 2003; Ketterings et al., 2001; Rahman et al., 2005), thus it is important to conduct further study on this topic.

5.2 Data and methods

5.2.1 Field Observation Data

This work used 1460 sampling plots allocated to 16 transects, and the size of each plot was approximately 225 m². In total, 13048 trees with diameters from 10 – 210 cm were measured and used to calculate basal area per hectare and stem volume per hectare using the allometric models adjusted for specific tree species (Tab. 5.1). Above ground biomass (AGB) was estimated subsequently using diameter at breast height (DBH) – biomass conversion model developed for low dipterocarp forests (Samalca, 2007).

The stem volume varied from 1.73 – 628.62 m³/ha and the mean volume was 156±92 m³/ha. Similar with stem volume, the AGB also showed highly variable values and the mean AGB was 167±94 ton/ha. These variations are common for natural forests especially those which are occupied by secondary and regenerating forests. Tree regenerating processes take place following the completion of forest harvesting, forest burning, and other types of forest disturbance. These processes which can continue for over 30 years are affected by various intrinsic and extrinsic aspects, e.g. anthropogenic factors, drought, disease, etc.

5.2.2 Images Acquisition and Preprocessing

Two sets of Landsat 7 ETM+ images with 30 meter spatial resolution were used. The first Landsat image was acquired on August 26, 2000 under hazy and cloud conditions, and the second image, acquired on May 31, 2003, showed clear atmospheric conditions with no apparent clouds. The satellite data were orthorectified into WGS 84 datum and projected on Zone 50N using Universal Transverse Mercator (UTM) projection. Preprocessing of ETM images were conducted for correcting the atmospheric and topographic effects to minimize the artifacts caused by the atmospheric attenuations, e.g. haze and irradiance scattering, and the terrain effects.

Tab. 5.1 Descriptions of sampling plots describing parameters of different forest physiognomies

	Shrub (Sh)	Riparian forest (RF)	Dense forest (DF)	Very dense forest (VDF)	Mature forest (MF)
Number of Stems (stems/ha)	219.2±125.7	290.6±104.8	224.5±110.0	319.3±168.5	262.2±281.8
DBH (cm)	25.7±9.4	26.7±4.9	34.1±9.7	32.7±8.5	54.1±24.3
Basal Area (m ² /ha)	8.9±6.0	11.8±5.1	12.3±6.2	16.6±8.1	18.2±11.8
Stem Volume (m ³ /ha)	92.4±69.5	121.5±60.3	142.5±81.3	189.4±97.7	221.1±144.4
Biomass (ton/ha)	105.4±72.2	139.0±63.1	152.4±82.6	200.4±99.7	234.2±154.7
Number of plots (n = 1460)	58	24	885	455	38
Descriptions	Mixture of pioneer species, low to medium tree size and shrubs, canopy cover < 50%, currently disturbed, shows noticeable marks of forest burning and clearing	Sparse forest dominated with slim and tall vegetation, canopy cover < 50%, located adjacent to the streams	Dense forest (canopy cover 50 - 70%), logged over < 10 years, located in flat and moderate slope	Very dense forest (canopy cover 70 - 80%), logged over between 10 - 20 years, located in moderate and highly steep regions	Advanced forest structure, closed canopy (over 80%), logged over > 20 years, located mostly in highly steep region

Moreover, calculation of vegetation indices required the surface reflectance rather than digital number (DN) values or top of atmosphere reflectance, thus the corrections on the images were required.

Atmospheric corrections were applied on the ETM data using dark object subtraction (DOS) method proposed by Chavez (1988). According to a study conducted by Song et al. (2001), different variations of DOS technique are available. We experienced the COST-DOS technique offered more preferable results with regard to the spectral responses of vegetated areas. Topographic corrections were implemented using C-Correction procedure assuming Lambertian effects on the earth surface (Riaño et al., 2003). Hereafter, we refer the satellite images to the corrected ETM data.

Digital Elevation Model (DEM) of the area was obtained from the Shuttle Radar Topography Mission (SRTM) data. The DEM originally 90 meter resolution was orthorectified with the ETM data and resampled using nearest neighborhood method into 30 meter spatial resolution to fit with the resolution of the ETM image. Slope angle and aspect were computed from the resampled DEM and applied as ancillary input for AGB and stand volume modeling.

5.2.3 Methods

Generally, this study approached the above ground biomass (AGB) and stem volume using RS data and synergy of GIS – field observation data (Fig. 5.1). To estimate AGB using field observation approach and GIS, we used a stem diameter (DBH) – AGB allometric equation developed for tropical lowland dipterocarp forest (Eq. 5.1). The following equation was generated by destructively measuring 40 sampling trees (Samalca, 2007).

$$AGB = \exp(-1.2495 + 2.3109 \times \ln(dbh)) \quad \text{Eq. 5.1}$$

The AGB estimated from Eq. 5.1 was superimposed over the modified GIS land cover map provided by the forest management unit to analyze the AGB density of particular land cover type. The land cover map was based on Landsat ETM image interpretation coupled with field surveying data.

The corrected Landsat ETM images were classified using Maximum likelihood method, and post-classification processing was carried out implementing majority analysis for removing minor spurious pixels within a large single class. In the majority analysis, we set up a

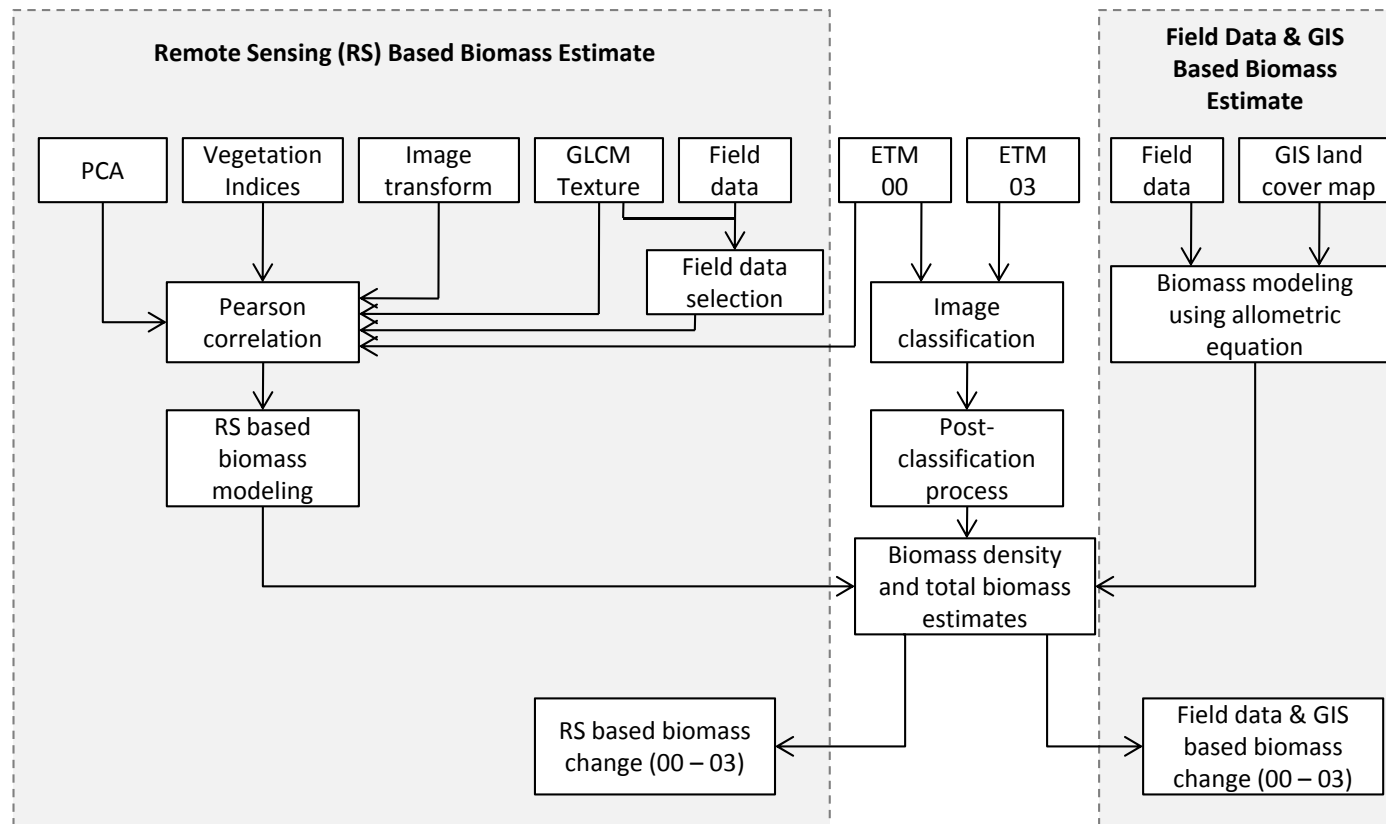


Fig. 5.1 Workflow of the study describes two main approaches for estimating the AGB, using remote sensing method (left shaded box) and combination of field data and GIS method (right shaded box). The middle part of the workflow (non shaded area) explains the classification procedure of multi-temporal ETM images (2000 and 2003) performed in this study

parameter of kernel size, using which the center pixel in the kernel was replaced with the class label that was dominant within this kernel. This process was iterated for the entire image resulting in a more homogenous classified image. Principal Component Analysis (PCA) bands were also estimated and used for the classification alternatively. The classification accuracy was assessed using confusion matrices and the associated kappa statistics.

To integrate the remote sensing data in estimating the AGB, the Landsat ETM data, vegetation indices (VI), simple ratio (SR), image transform data (i.e. VIS123, ALBEDO, MID57), tasseled cap (TC), three bands of principal component analysis (PCA), GLCM texture features, and slope and aspect from the DEM data were statistically correlated with the biomass data following the Pearson correlation test procedure. The biomass data was carefully selected using GLCM mean texture feature as the reference, as we obtained this texture feature had the highest correlation coefficient compared to other RS data. The modeling of the AGB and stem volume equations was conducted using SPSS version 11.5 software applying a stepwise multi-linear regression method. The modeling used a subset of sample data, and was validated with the complete dataset. Biomass density and total biomass of each land cover class was predicted overlaying the AGB estimate with the land cover maps of 2000 and 2003. Subsequently, the total biomass change during this period was calculated. Dynamics of the estimated forest properties assessed from RS and combination of GIS – field observation approaches, and substantial correlation between GLCM mean texture and the AGB are discussed.

5.2.4 Vegetation Indices Generation and Land Cover Classification

Various vegetation indices may be computed from the satellite data. These vegetation indices were proposed for different applications, such as soil moisture, vegetation monitoring, mineral deposits mapping, etc (Jensen, 1996). Vegetation indices generated from certain satellite image bands are sensitive to characterize green vegetation/forested regions from other objects on the ground. In vegetated regions, the cells in plant leaves are very effective scatterers of light because of the high contrast in the index of refraction between the water-rich cell contents and the intercellular air spaces. Vegetation is very dark in the visible bands (400-700 nm) because of the high absorption of pigments in leaves (chlorophyll, protochlorophyll, xanthophyll, etc.). There is a slight increase in reflectivity around 550 nm (visible green band) because the pigments are least absorptive in this range. In the spectral range of 700-1300 nm plants are very bright because this is a spectral non-man's land between the electronic transitions, providing absorption in the visible and molecular vibrations that absorb in longer wavelengths. There is no strong absorption in

this spectral range, but the plant scatters strongly. From 1300 nm to about 2500 nm vegetation is relatively dark, primarily because of the absorption by leaf water. Cellulose, lignin, and other plant materials are also absorbed in this spectral range (Lillesand and Kiefer, 1994). This study, moreover, demonstrated the utility of vegetation indices, especially those proposed for vegetation monitoring, for estimating the AGB and stand volume (Tab. 5.2). Besides those indices explained above, we also computed three bands of principal component analysis (PCA) and three bands of tasseled cap (TC), i.e. brightness (TC1), greenness (TC2) and wetness (TC3). Another attempt to include more Landsat ETM features was to calculate the Gray Level Co-Occurrence Matrix (GLCM) texture features. Eight GLCM texture computed using second derivatives of mean (GLCM_MEAN), variance (GLCM_VAR), homogeneity (GLCM_HOMO), contrast (GLCM_CONT), dissimilarity (GLCM DISS), entropy (GLCM_ENTR), second moment (GLCM_SECM), and correlation (GLCM_CORR) were generated. We analyzed the variance matrix of Landsat ETM bands and found substantial variance of forested lands from Landsat ETM band 5 (Wijaya et al., 2008b). This band was ultimately selected for generating the texture features using 5×5 moving window (cf. chapter 3 for moving window size selection procedures). The texture layers were calculated to each direction with single shifting pixel and were quantified into a 64 gray levels.

5.3 Results

5.3.1 Biomass Mapping using Field Data and GIS

The estimated biomass, stem volume and count of stems were plotted against DBH classes (Fig. 5.2), and the result show that small and medium tree diameters (10 - 60 cm), although contributing to large number of trees, represented small amounts of stem volume and biomass. These DBH's were dominant for young- and regenerating forests, which were mostly occupied by small and fast growing pioneer species, e.g. *Macaranga* sp., complemented with medium size of non-pioneer species, e.g. *Aglaia*, *Knepa* and *Artocarpus*. Old secondary- and mature forests on the other hand, were characterized by non-pioneer species from medium to large tree size, e.g. *Shorea*, *Dipterocarpus*, *Vatica* and *Euphorbiaceae* species. The pioneer species and small light demanding species disappeared during the regenerating process due to natural thinning effects caused by species competition in pursuing limited nutrient contents and light intensities. Therefore, the old secondary- and mature forests contributed to higher AGB and stem volume.

Tab. 5.2 Various simple ratios, different NDVIs, complex vegetation indices and image transform layers computed in this study to generate the biomass and stem volume equations

Index under study	Formula	Description	Reference
Simple band ratios			
ETM 4/3	$ETM4/ETM3$	Measure of green vegetation and chlorophyll absorption bands (ETM _{mid} : 660nm; 830nm)	(Rouse et al., 1973)
ETM 5/3	$ETM5/ETM3$	Modified simple ratio (ETM _{mid} : 660nm; 1650nm)	(Lu et al., 2004)
ETM 5/4	$ETM5/ETM4$	Modified simple ratio (ETM _{mid} : 830nm; 1650nm)	(Lu et al., 2004)
ETM 5/7	$ETM5/ETM7$	Modified simple ratio (ETM _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)
ETM 7/3	$ETM7/ETM3$	Modified simple ratio (ETM _{mid} : 660nm; 2215nm)	(Lu et al., 2004)
Traditional vegetation indices			
NDVI	$(ETM\ 4 - ETM\ 3)/(ETM\ 4 + ETM\ 3)$	Measure of green vegetation cover (ETM _{mid} : 660nm; 830nm)	(Rouse et al., 1973)
ND53	$(ETM\ 5 - ETM\ 3)/(ETM\ 5 + ETM\ 3)$	Modified vegetation indices (ETM _{mid} : 660nm; 1650nm)	(Lu et al., 2004)
ND54	$(ETM\ 5 - ETM\ 4)/(ETM\ 5 + ETM\ 4)$	Modified vegetation indices (ETM _{mid} : 830nm; 1650nm)	(Lu et al., 2004)
ND57	$(ETM\ 5 - ETM\ 7)/(ETM\ 5 + ETM\ 7)$	Modified vegetation indices (ETM _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)

ND32	$(ETM\ 3 - ETM\ 2)/(ETM\ 3 + ETM\ 2)$	Modified vegetation indices (ETM+ _{mid} : (Lu et al., 2004) 560nm; 660nm)	
Complex vegetation indices			
ARVI	$(NIR + 2RED + BLUE)/(NIR + 2RED - BLUE)$	Enhancement of NDVI that is relatively (Kaufman and Tanre, 1996) resistant to atmospheric factors (ETM+ _{mid} : 485nm; 660nm; 830nm)	
EVI	$2.5 \times (NIR - RED)/(NIR - 6RED - 7.5BLUE + 1)$	Reduce the atmospheric influence and (Huete et al., 1997) optimize the vegetation signal (ETM+ _{mid} : 485nm; 660nm; 830nm)	
SAVI	$(NIR - RED) \times (1 + L)/(NIR + RED + L)$	Modified green vegetation index with an (Huete, 1988) adjustment factor (ETM+ _{mid} : 660nm; 830nm)	
MSAVI2	$((2NIR+1)-\sqrt{(2NIR+1)^2-8(NIR-2RED)})/2$	Measure of vegetation that is less (Qi et al., 1994) sensitive to atmosphere and soils (ETM+ _{mid} : 660nm; 830nm)	
GEMI	$\varepsilon(1-0.25\varepsilon)-(RED-0.125)/(1-RED)$ where $\varepsilon = (2(NIR^2-RED^2)+1.5NIR+0.5RED)/(NIR+RED+0.5)$	Global environmental monitoring index (Pinty and Verstraete, 1991) that is insensitive to atmosphere (ETM+ _{mid} : 660nm; 830nm)	
Image transform indices			
VIS123	$ETM1 + ETM2 + ETM3$	(ETM+ _{mid} : 485nm; 560nm; 660nm)	(Lu et al., 2004)
MID57	$ETM5 + ETM7$	(ETM+ _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)
ALBEDO	$ETM1 + ETM2 + ETM3 + ETM4 + ETM5 + ETM7$	(ETM+ _{mid} : 485nm; 560nm; 660nm; 830nm; 1650nm; 2215nm)	(Lu et al., 2004)

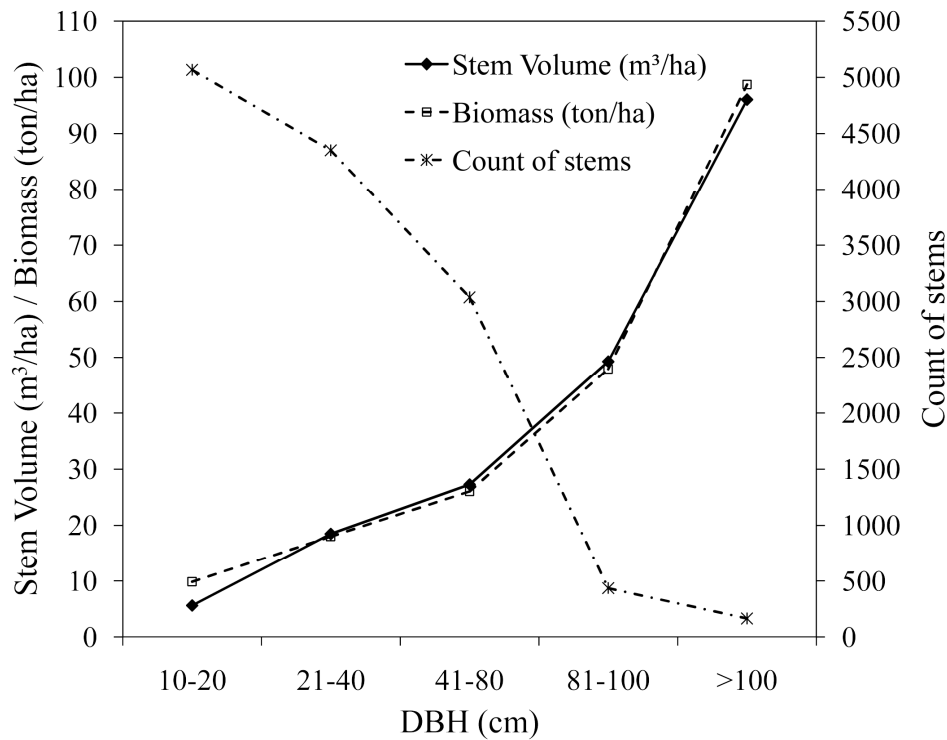


Fig. 5.2 Plot of Stem volume/biomass and number of stems/ha vs. tree diameter (DBH)

The AGB was exponentially increased with DBH following the forest regenerating processes. This is because the biomass has an exponential relationship with DBH, and stem volume is basically a square function of DBH. Assuming other conditions are constant, stem volume is linearly related to AGB. Besides that, biomass of a single tree is equal to the product of the wood density and the volume (Ketterings et al., 2001). Many large trees in our study area comprise of hardwood trees, such as teak (*Tectona grandis*), mahogany (*Swietenia* sp.), ebony (*Diospyros* sp.), keruing (*Dipterocarpus* sp.), and meranti (*Shorea* sp.). The hardwoods are mostly broad-leaved, and in the tropics and subtropics these trees are usually evergreen. On average, hardwood has higher wood density and hardness than softwood, although there is an enormous variation in actual wood hardness in both groups, with the range in density in hardwoods completely including that of softwoods.

Based on the interpretation of the Landsat image and DEM data in 2001, eleven land cover classes were identified in the concession area. The forest management unit digitized GIS land cover map and used it as the reference in managing the concession area (Fig. 5.3). Mature forest was defined as an old forest comprising large growing trees and some patches of primary forest. Very dense forests was explained as old secondary forest, which were logged >20 years ago and comprised of more large trees rather than the regenerating ones. Dense forests were described as current regenerating

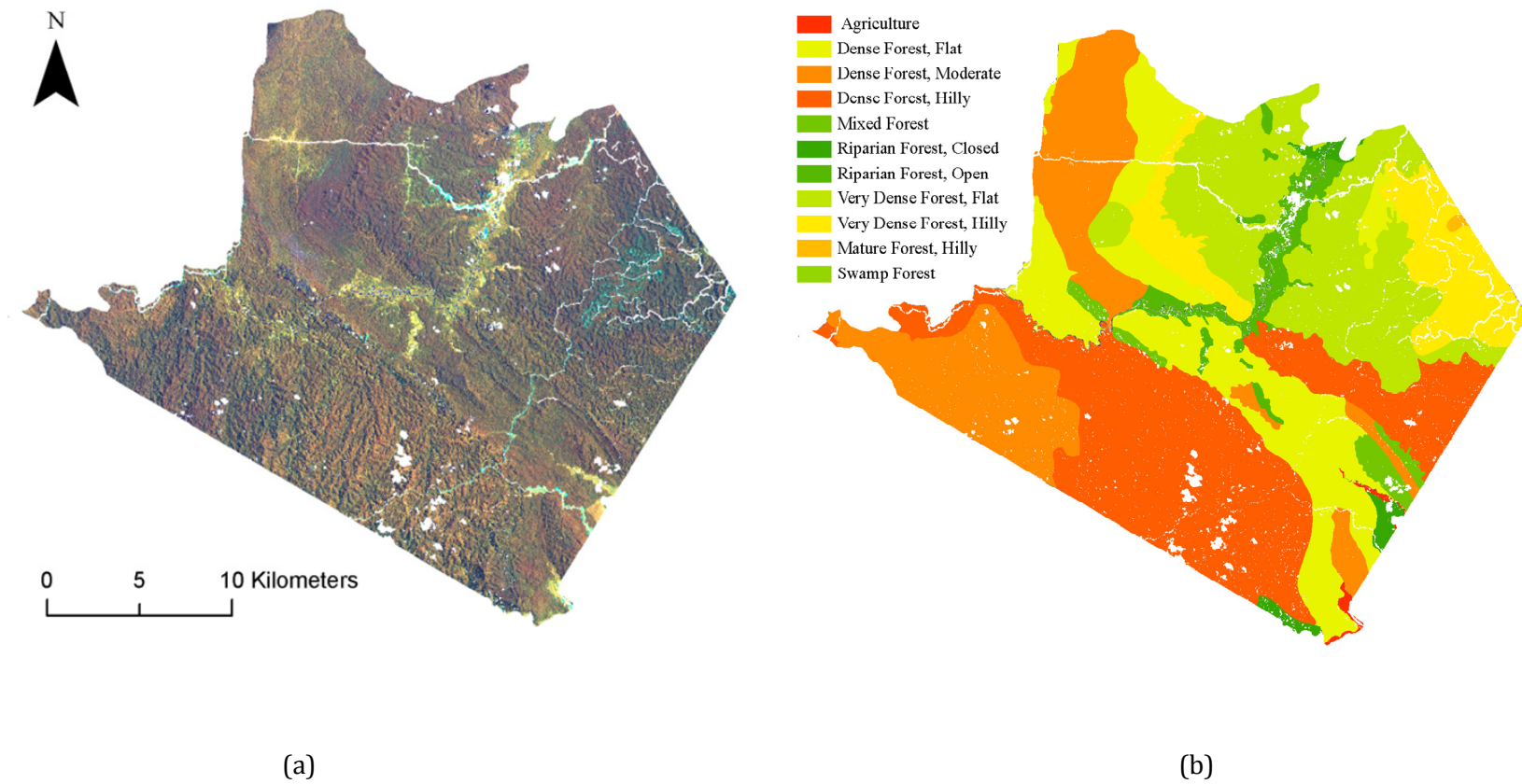


Fig. 5.3 Landsat ETM of the study area with ETM bands 453 as RGB combination (a), and modified GIS land cover map of 2001 (b)

forests with the age of <20 years old. Riparian forests were situated along the main rivers that flow over the study area from NE – SW to west directions.

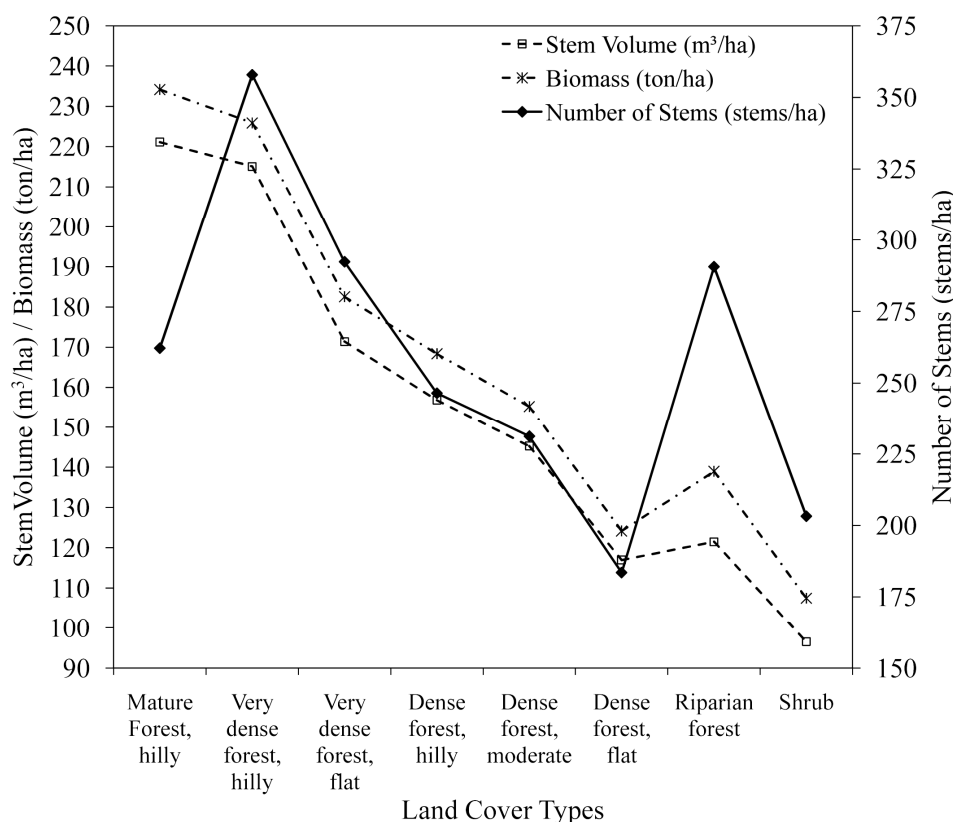


Fig. 5.4 Above ground biomass (AGB) and stem volume of each land cover type, sorted with the most advanced vegetation structures, i.e. mature forest-hilly, to the least complex structures, i.e. shrubs

Using GIS land cover map as a reference, the stem volume, AGB and number of stems per hectare were plotted (Fig. 5.4). The mean AGB of mature forest was 234.2 ton/ha, almost double than in dense forests (149.8 ton/ha). The stem number in mature forest was extremely high, mainly due to the presence of undisturbed forests within this class which mostly situated in the protection forest. Riparian forest on the other hand, represented higher stem number but contributed to low biomass. This was because most riparian forests were characterized with tall and very slim trees, thus contributed to less biomass. The presence of riparian forest was usually mixed with shrubs which may become a problem for RS data to classify both forest classes. Similarly, a number of shrubs was also found along the rivers indicating the presence of former slash and burn practice and opening of agricultural lands.

During field work, we observed more disturbance on forests occurred with declining slope. These disturbances were mainly caused by anthropogenic factors, such as illegal forest harvesting, forest burning and opening of agricultural farms (Wijaya, 2006). This condition

was similar with the finding of this study which predicted lower AGB and stem volume in less steep areas (Fig. 5.4). For example, the biomass density in very dense forest-hilly (225.9 ton/ha) was significantly higher than in very dense forest-flat (182.5 ton/ha), while the AGB in dense forests under moderate (155.2 ton/ha) and under hilly terrains (168.4 ton/ha) was slightly different.

Our ground observation also suggested that vegetation complexity and canopy cover in mature forests and very dense forests were similar, given the fact that the estimated AGB in mature forests (234.2 ton/ha) was slightly higher than in very dense forests (204.2 ton/ha). As mentioned earlier, mature forests indicate the presence of pristine forests which mostly are undisturbed, whereas very dense forests were assumed as old secondary forests that were harvested more than 20 years ago. The similarity in vegetation structures for both forests occurred because these forests were selectively logged over, depending on tree species (i.e. commercial timber) and size (i.e. DBH > 50 cm), and the gaps of forest canopy were rapidly recovered just one year after the completion of forest harvesting. However, the density of large trees (DBH > 80 cm) in secondary forests (i.e. very dense forests) was not as high as in primary forests (i.e. mature forests), so that lower biomass was found in the secondary forests even after 20 years of forest harvesting, as indicated by number of stems per hectare in both forest regimes (Fig. 5.4).

5.3.2 Prediction and Dynamics of Biomass and Stem Volume using Remote Sensing

5.3.2.1 Model generation

Analysis of Pearson correlation test showed that GLCM mean texture explained the stem volume ($r = -0.669$) and AGB ($r = -0.544$) better than other remote sensing data, including Landsat ETM band 4, 5 and 7, NDVI, SAVI, and PC1, which were usually among the 'favorite bands' used for vegetation assessment (Tab. 5.3).

This high correlation between GLCM mean texture and AGB and stem volume was probably due to the smoothing effects of the texture feature calculating second derivatives mean values of particular pixels based on the values of neighboring pixels. The utility of GLCM texture features was useful to remove the shadow effects of broadleaf and/or large trees and shows higher correlation between the GLCM entropy and AGB in mature tropical forests of the Amazon (Lu, 2005).

Tab. 5.3 Correlations between Remote Sensing Data, Stem Volume and Above Ground Biomass (AGB)

	Stem Volume	AGB		Stem Volume	AGB
B1	-.250(**)	-.246(**)	SAVI	-.101	-.076
B2	-.278(**)	-.275(**)	MSAVI2	-.075	-.052
B3	-.211(**)	-.194(**)	GEMI	.368(**)	.313(**)
B4	-.395(**)	-.336(**)	VIS123	-.276(**)	-.267(**)
B5	-.418(**)	-.375(**)	MID57	-.425(**)	-.382(**)
B7	-.390(**)	-.349(**)	ALBEDO	-.443(**)	-.394(**)
ELEV	-.009	-.160(**)	PC1	-.426(**)	-.383(**)
SLOPE	.082	.082	PC2	.399(**)	.340(**)
SR	-.137(*)	-.109(*)	PC3	-.077	-.085
SR53	-.177(**)	-.160(**)	TC1_BR	-.427(**)	-.374(**)
SR54	-.082	-.093	TC2_GR	-.300(**)	-.241(**)
SR57	-.037	-.037	TC3_WE	.380(**)	.338(**)
SR73	-.147(**)	-.130(*)	GLCM_MEAN	-.669(**)	-.544(**)
NDVI	-.085	-.063	GLCM_VAR	-.067	-.035
ND53	-.129(*)	-.111(*)	GLCM_HOMO	.081	.078
ND54	-.084	-.095	GLCM_CONT	-.100	-.063
ND57	-.034	-.032	GLCM DISS	-.099	-.085
ND32	.061	.083	GLCM_ENTR	-.011	-.016
ARVI	-.099	-.073	GLCM_SECM	.011	.027
EVI	.014	-.002	GLCM_CORR	-.020	-.016

That study explained the combination of texture features and Landsat ETM spectral data could improve the predictive ability of multi-linear regression method in estimating the AGB. Another study carried out in the regenerating tropical forest of Brazilian Amazon described that GLCM contrast improved the correlation between radar backscatter and the AGB (Kuplich et al., 2005). The study area in fact, was a secondary forest that mostly classified as moderate – late regenerating forest stages and mature forest (as shown in Fig.

5.5). With relatively complex vegetation structure, the GLCM texture features were more sensitive to AGB than Landsat ETM spectral data and vegetation indices.

Providing the highest correlation coefficients with the AGB and stem volume, the GLCM mean texture was ultimately used as a basis for sample data selection, resulting in the subset data exhibited in Tab. 5.4. The data selection was actually a process to remove the extreme values from the complete dataset, and hence, reduced standard deviation of the subset data. Comparison between the subset and complete datasets found similar mean AGB and stem volume, and spatial distribution of the data was also similar with the complete dataset, as only the data within $\pm 1SD$ were selected. Assuming there was no change on the data distribution, the subset data was used to build the remote sensing-based AGB and stem volume linear equations. Landsat ETM multispectral bands (ETM Bands 1-5, and 7), SR53, SR73, GEMI, VIS123, MID57, ALBEDO, PC1, TC1, TC2, TC3, and GLCM mean were significantly correlated with AGB and stem volume ($\rho < 0.05$), although the correlation coefficients on average were relatively low ($r < 0.5$) (Tab. 5.3).

Tab. 5.4 Comparison of stem volume and AGB from complete and selected datasets

	Complete Dataset		Subset Data	
	Stem Volume (m ³ /ha)	AGB (ton/ha)	Stem Volume (m ³ /ha)	AGB (ton/ha)
Mean	156.79	167.36	156.60	166.82
Min	1.73	4.69	59.95	60.85
Max	628.62	663.35	221.97	234.03
SD	92.15	94.16	24.69	27.12
%SD	59%	56%	16%	16%
N	1460		388	

Using Stepwise method, these data were iteratively selected to model the stem volume (StVol), and the linear equation model was generated (SEE=18.4, F=34.719, $\rho < 0.05$).

$$StVol = (9.703 \times B4) + (11.910 \times B5) + (8.51 \times B7) + (0.001 \times GEMI) - (22.444 \times ALBEDO) + (4214.699 \times PC1) - (254.412 \times TC3_WE) - (15.595 \times GLC_MEAN) + 1192.511 \quad \text{Eq. 5.2}$$

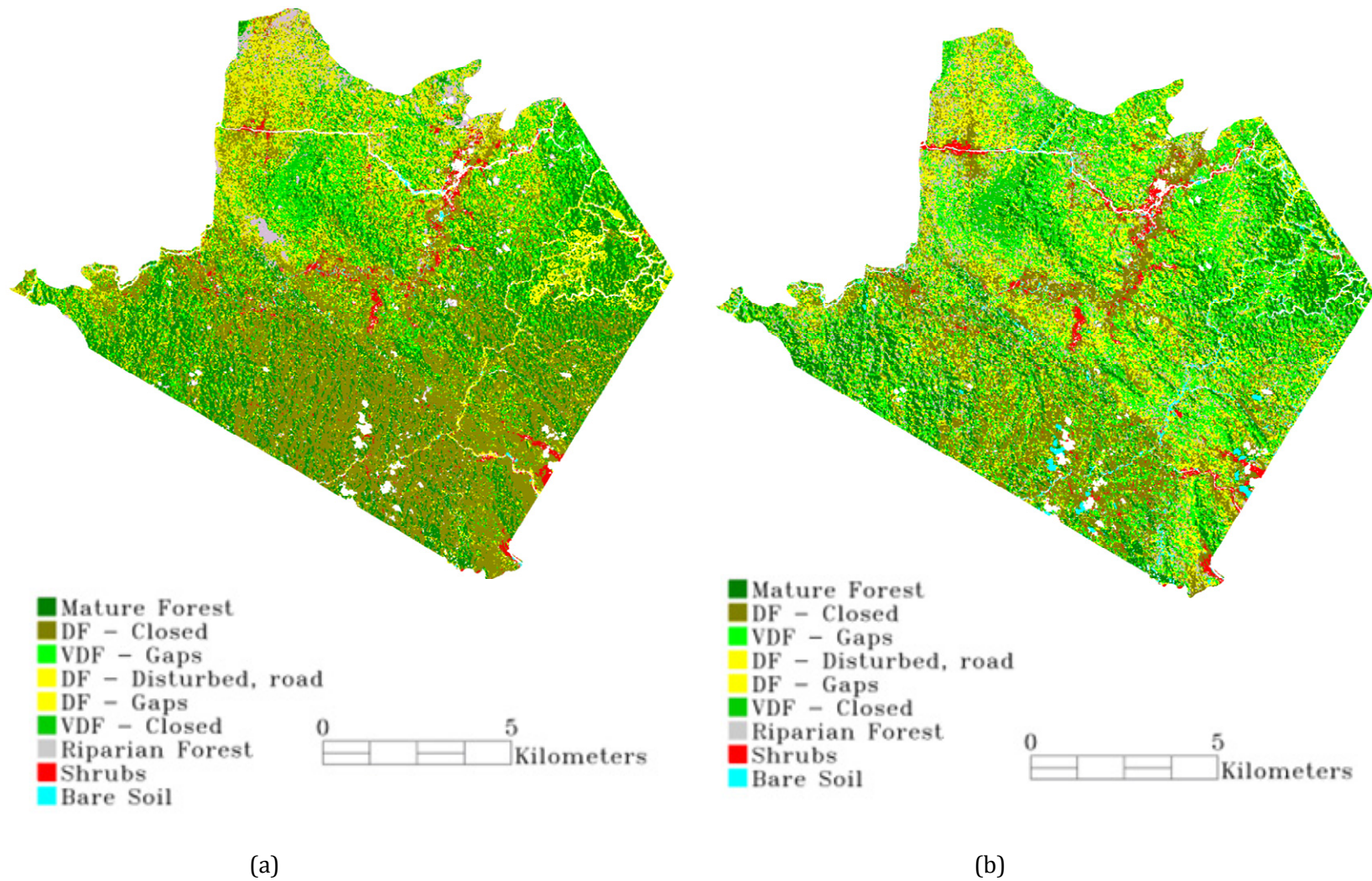


Fig. 5.5 Classified Landsat ETM images of year 2000 (a) and 2003 (b) showing mature forest, very dense forests, dense forests, riparian forest, shrubs and bare soil. The bare soil class was masked out from the classification prior to the estimates of AGB density and stem volume of each land cover type

Similar to stem volume, the AGB was estimated using combination of the RS data for predicting the biomass linear equation ($SEE=22.7$, $F=21.44$, $\rho < 0.05$).

$$AGB = (6.569 \times B4) - (14.198 \times B5) - (9.366 \times B7) - (14.784 \times ALBEDO) + (3430.451 \times PC1) - (2647.087 \times TC3_WE) - (12.991 \times GLC_MEAN) + 1029.644 \quad \text{Eq. 5.3}$$

Applying Equations (2) and (3), we estimated 157.8 ± 16.12 m³/ha of stem volume and 168.06 ± 14.57 ton/ha of AGB over the study area. These estimates were similar with those predicted from the field observation data obtaining 156.79 ± 92.15 m³/ha and 167.36 ± 94.16 ton/ha of stem volume and AGB, respectively (Tab. 5.4).

5.3.2.2 Land cover classification

The accuracy of classification results was assessed using confusion matrices and Kappa Statistics (Tab. 5.5), and found the classification using Landsat ETM image and processed using majority analysis had better accuracy (OA00 = 82.8%, OA03 = 85.1%) than the use of PCA bands (OA00 = 75.9%, OA03 = 80.8%) or Landsat ETM data without post-classification process (OA00 = 77.9%, OA03 = 81.9%).

Tab. 5.5 Classification Accuracy of Landsat ETM 2000 and 2003

	ETM 2000		ETM 2003	
	OA ₀₀ (%)	Kappa	OA ₀₃ (%)	Kappa
ETM image (Band 1-5,7)	77.9	0.75	81.9	0.79
PCA Bands (PC 1-3)	75.9	0.73	80.8	0.78
ETM image, majority analysis	82.8	0.80	85.1	0.83

Majority analysis was basically an attempt to remove minor spurious pixels surrounded within a large single class using a kernel matrix. The analysis resulted in more homogenous classification map, which had higher accuracies and better visualization characteristics.

Based on the Landsat ETM data, nine land cover classes, namely mature forest, very dense forest classes (VDF-closed, VDF-gaps), dense forest classes (DF-closed, DF-gaps, DF-disturbed), riparian forest (RF), shrubs and bare soil were classified. The classification map showed noticeable marks of forest degradation and deforestation from 2000 to 2003 (Fig. 5.5). Southern part of the study area, which were dominated by very dense forests in 2000 were mostly converted into dense forest in 2003, indicating prominent forest degradation. The expansion of road networks and slash and burn practice for the opening of new

agriculture lands were the major problems compromising the sustainability of forest management over this forest region.

5.3.3 Comparison of AGB and stem volume estimates

We have so far two land cover maps, namely the GIS land cover map and land cover map of Landsat ETM data classification. Unfortunately, both maps have different number of classes and class descriptions. There were eleven classes and nine land cover types identified in the GIS land cover map and the classification of Landsat ETM data, respectively. To compare the biomass density and standing stocks estimated from RS and GIS – field observation based approaches, the incompatible class labels were excluded or aggregated following general classification rule. The incompatible land cover classes, i.e. agriculture, mixed forest and swamp forest classes were excluded from the GIS land cover map. The remaining classes were aggregated resulting in five final classes for both land cover maps, namely mature forest, very dense forest, dense forest, riparian forest and shrubs. Using the aggregated land cover classes, the assessments of biomass and stem volume changes from 2000 to 2003 were conducted.

Tab. 5.6 Comparison of AGB and stem volume estimates for particular land cover type

	Remote Sensing Estimate		GIS and Field Data Estimate	
	Stem Volume	AGB	Stem Volume	AGB
	(m ³ /ha)	(ton/ha)	(m ³ /ha)	(ton/ha)
Mature forest	171.8	180.8	271.1	287.9
Very dense forest	161.5	171.0	193.3	204.2
Dense forest	152.8	163.4	140.3	149.8
Riparian forest	146.1	154.1	121.5	139.0
Shrubs	105.2	118.7	96.5	107.5

An attempt using GIS and field data obtained higher AGB and stem volume in mature- and very dense forests classes than did the remote sensing approach. For dense forest, riparian forest and shrubs, the AGB and stem volume were higher when the RS data was applied (Tab. 5.6). In general, both approaches lead to similar conclusion, as the regenerating stage becomes more advanced, the more AGB and standing stocks were found in the study area. This study considered shrubs as the least complex vegetation structure, representing the

earliest regenerating stage. In contrast, mature forest was associated with the most advanced vegetation structure.

5.3.4 Dynamics of biomass abundance

The land cover map of classified Landsat ETM data was used as the reference for analyzing the total areas of each land cover type and biomass change from 2000 to 2003. Of about 4,200 ha of mature forests in 2000 were converted into other land cover types in 2003 showing forest degradation within this particular forest.

Mature forests are important for forest ecosystem, as these forests represent the most complex vegetation structure and indicate the presence of undisturbed forests. Very dense forests increased from 8,859 ha to 16,865 ha (Tab. 5.7). This was probably due to the degradation of mature forests, or due to the growth of dense forest into a more complex structure reducing the area of this particular class from 35,563 ha (2000) to 27,624 ha (2003). The riparian forest areas increased up to 4,962 ha in 2003, this implied the excessive extension of shrubs into respective forest class. The bare soil class was none of our interest, therefore excluded prior to the biomass change assessment.

Tab. 5.7 Percentage of Land cover change from 2000 to 2003 based on Landsat ETM data classification (percentage is shown in brackets)

	Land Cover 2000 (ha) / (%)	Land Cover 2003 (ha) / (%)	Difference (2003 – 2000) (ha) / (%)
Mature forest	15,297 (24.4%)	11,094 (17.7%)	-4,202 (-6.7%)
Very dense forest	8,859 (14.2%)	16,865 (27.0%)	8,006 (12.8%)
Dense forest	35,563 (56.8%)	27,624 (44.2%)	-7,939 (-12.7%)
Riparian forest	1,550 (2.5%)	4,962 (7.9%)	3,413 (5.5%)
Shrub	1,150 (1.8%)	1,134 (1.8%)	-15 (0.0%)
Bare Soil	151 (0.2%)	888 (1.4%)	737 (1.2%)
Total classified area	62,568 (100%)	62,568 (100%)	

Calculating the sum products of AGB and stem volume (Tab. 5.6) and total forest areas in 2000 and 2003 (Tab. 5.7), the changes on biomass and standing stocks over the study area were obtained (Tab. 5.8). The AGB in mature forest decreased by 25% from 2.77 Gt in 2000 to 2.0 Gt in 2003 estimated from RS data. Similarly, the GIS – field data assessed

Tab. 5.8 Dynamics of Forest Biomass (AGB) and Stem Volume (Vol) from 2000 to 2003

Remote Sensing data estimate										
	2000				2003				Differences (2003 – 2000)	
	Vol (m³)	(%)	AGB (Gton)	(%)	Vol (m³)	(%)	AGB (Gton)	(%)	Vol (m³)	AGB (Gton)
Mature forest	2,627,681	26.7%	2.766	26.4%	1,905,789	19.7%	2.006	19.5%	-721,893	-0.760
Very dense forest	1,430,454	14.5%	1.515	14.5%	2,723,228	28.1%	2.884	28.0%	1,292,774	1.369
Dense forest	5,435,006	55.2%	5.812	55.5%	4,221,747	43.5%	4.515	43.8%	-1,213,260	-1.298
Riparian forest	226,391	2.3%	0.239	2.3%	725,013	7.5%	0.765	7.4%	498,621	0.526
Shrub	120,894	1.2%	0.136	1.3%	119,295	1.2%	0.135	1.3%	-1,599	-0.002
Total	9,840,427	100%	10.469	100.0%	9,695,071	100.0%	10.304	100.0%	-145,356	-0.164
GIS and field data estimate										
	2000				2003				Differences (2003 – 2000)	
	Vol (m³)	(%)	AGB (Gton)	(%)	Vol (m³)	(%)	AGB (Gton)	(%)	Vol (m³)	AGB (Gton)
Mature forest	4,147,107	37.2%	4.403	37.1%	3,007,788	27.7%	3.194	27.6%	-1,139,319	-1.210
Very dense forest	1,712,045	15.4%	1.809	15.2%	3,259,308	30.0%	3.444	29.7%	1,547,262	1.635
Dense forest	4,987,952	44.7%	5.329	44.9%	3,874,489	35.7%	4.139	35.7%	-1,113,463	-1.189
Riparian forest	188,261	1.7%	0.215	1.8%	602,902	5.6%	0.690	6.0%	414,641	0.474
Shrub	110,976	1.0%	0.124	1.0%	109,508	1.0%	0.122	1.1%	-1,468	-0.002
Total	11,146,342	100.0%	11.880	100.0%	10,853,995	100.0%	11.588	100.0%	-292,347	-0.292

lower biomass in this particular forest with greater magnitude. Both approaches found an increased AGB in very dense forest and riparian forest. The dense forest class, representing of more than 56% of forested lands, contributed to over 55% of total biomass in 2000 estimated using RS data. In overall, there was a slightly declining trend in total biomass (range from -0.164 Gt to -0.292 Gt) from 2000 – 2003. This indicates continuous degradation and deforestation within the forest region and consequently reduced the total abundance of biomass and the volume of standing stocks.

Carbon accumulation over this period definitely was reduced, and more carbon was released into the atmosphere. Remote sensing approach in general calculated lower biomass abundance and stem volume than those from GIS – field data. The earlier approach predicted 10.45 Gt and 10.3 Gt of total biomasses in 2000 and 2003, while the later estimated 11.9 Gt and 11.6 Gt of total biomasses, respectively.

5.4 Discussion

5.4.1 Prediction Results Assessment

This study successfully predicted the above ground biomass (AGB) and stem volume over a tropical forest region using remote sensing and GIS – field data approaches. Estimation of stem volume is important for mapping of standing stock and for forest inventory purpose, as it provides initial prediction on timber amount that could be commercially harvested. The biomass on the other hand is important for indicating carbon accumulation in a forest region over time, and information on total AGB estimated for each land cover type is useful to assess how different regeneration stages could have an effect on the forest as a source of carbon sink.

Remote sensing based estimates have potential to predict the dynamics of forest biomass and stem volume over large forest region with less efforts, time and cost than field based estimate. However, the accuracy of the estimates is somehow questionable, as it depends on the quality of remote sensing data and its relationship with field observation data being modeled.

Several attempts to estimate AGB from remote sensing data found high uncertainties which were around 30 – 40% (Sales et al., 2007). This study confirmed this high error estimate in assessing the AGB using RS data and found slightly lower error estimate (Tab. 5.9), and the result might be used as an initial prediction of AGB over the study area. To elevate the estimate precision, correlation analysis between the RS data and biomass can be separately implemented for different land cover, and it should be considered for further study.

Tab. 5.9 Above ground biomass estimates computed in this study using different allometric equations developed for tropical forest environment

	AGB _{GIS} (This study)	AGB _{RS} (This study)	FAO Model (1997)	Brown & Lugo (1992)	Ketterings, et.al. (2001)
AGB Estimate (ton/ha)	167.4	166.8	164	155.7	88.46
SD (ton/ha)	94.2	27.1	91.8	94.5	47.5

An attempt to estimate the AGB using remote sensing (RS) tends to underestimate the result due to the saturation of the Landsat ETM spectral values and vegetation indices. The RS data saturated at higher AGB and stem volume, reducing the coefficient correlation with the measured forest properties. In order to reduce the saturation problem, we masked the extreme values out from RS data to get better correlation with the forest properties under study. The present study as well as previous studies confirmed that reflectance of Landsat data and NDVI were saturated at higher biomass density (Lu, 2005; Steininger, 2000). Several underlying factors may cause this problem, namely the size of sampling plot that was not designed to be related to spaceborne data, or the saturation from dense leaf canopies that restricts the AGB estimates into a low level when passive sensors, such as Landsat ETM, are used (Anaya et al., 2009). Nevertheless, the utility of moderate resolution of satellite data, such as Landsat ETM image, is the only alternative to predict the AGB and stem volume in this particular forest due to the lack of active sensors, e.g. SAR and Lidar, or high spatial resolution satellite imagery, e.g. IKONOS and Quickbird.

The biomass estimates of this study were compared with those computed using another allometric model generated with destructive sampling and developed for similar forest environment. Assuming the similarity of forest structure and vegetation compositions, those models were implemented in this study for estimating the AGB using available sample dataset. Our estimates (AGB_{GIS} and AGB_{RS}) were similar with the results of FAO model (Brown, 1997) and Brown and Lugo study (Brown and Lugo, 1992). However, Ketterings model (Ketterings et al., 2001) estimated significantly lower biomass than did other models (Tab. 5.9). This was probably due to the forest composition in Sumatra, the site where this particular model was developed, did not represent the forest in the Labanan, although both forests were geographically located in one country. The AGB models developed for general tropical forest (Brown, 1997; Brown and Lugo, 1992) are more suitable for our study site. The Brown & Lugo model (Brown and Lugo, 1992) was generated collecting tree sample

from Brazil, Cambodia and Indonesia. Similarly, the FAO model (Brown, 1997) was developed for tropical moist forest environment in general.

5.4.2 Relationship between GLCM Mean Texture, Land Cover, and Forest Biomass

We found texture features derived from the Grey Level Co-occurrence Matrix (GLCM) mean texture had a strong correlation with the AGB and stem volume (Tab. 5.3). To study the capability of mean texture feature in discriminating the AGB of particular land cover type, the GLCM mean texture, AGB estimate, and land cover type were plotted showing that moderate- and flat terrain-dense forests represented higher texture values compared to that of very dense forest classes (Fig. 5.6).

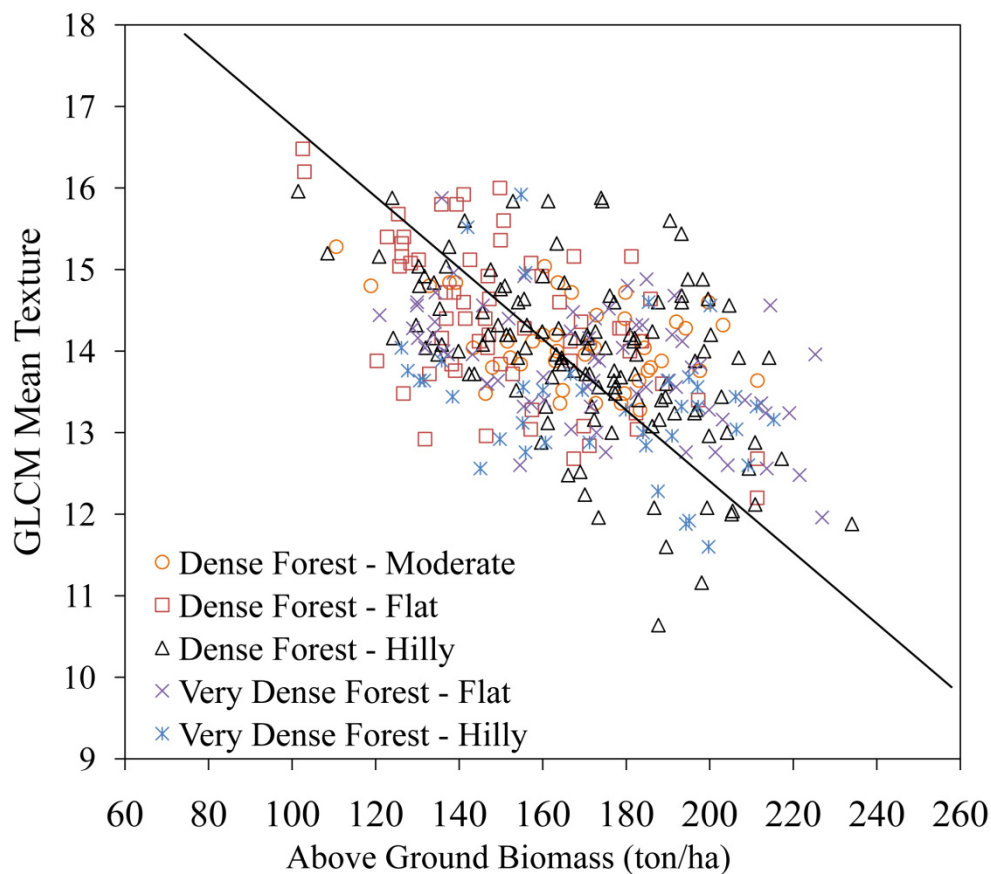


Fig. 5.6 Distribution of GLCM Mean Texture of Different Land Cover Type

The GLCM mean texture values of dense forest-hilly class have large interval and highly overlaps with other forests texture values. During field observation we found many similarities between very dense- and mature forests, and to differentiate these forest classes is sometimes problematic, especially those located in moderate and steep slope. Within these forest classes, we found numbers of moderate trees (dbh>50 cm) configured with a

very small gap of canopy opening. Calculating mean texture features, the shadow effects from tree canopies was removed, but the limited ability of Landsat ETM data in penetrating through the forest canopies created problems for characterizing each land cover biomass using individual texture data.

5.5 Conclusions

The assessment of above ground biomass (AGB) and stem volume was presented in this study implementing RS data and GIS – field data approach. The Landsat ETM data, vegetation indices, image transform layers, simple ratio, PCA, tasseled caps bands, GLCM texture features and DEM were generated and correlated with the AGB and stem volume. We found the GLCM mean texture had higher coefficient correlation than other RS data, but was difficult for discriminating the biomass of each land cover type due to the limitation of Landsat ETM data. Based on selected dataset, the linear equation models of AGB and stem volume were predicted. On average, 158 ± 16 m³/ha of stem volume and 168 ± 15 ton/ha of AGB were estimated using RS approach. Based upon the field observation data, 157 ± 92 m³/ha and 167 ± 94 ton/ha of stem volume and AGB were predicted, respectively. The dynamics of biomass abundance from 2000 to 2003 were assessed using classified Landsat ETM data. In general, there was a declining trend of total biomass over this period. Remote sensing approach estimated lower biomass abundance than did the GIS and field data. The earlier approach predicted 10.47 Gt and 10.3 Gt of total biomasses in 2000 and 2003, while the later estimated 11.9 Gt and 11.6 Gt of total biomasses, respectively.

Chapter 6

Retrieval of Forest Attributes using Bi-Temporal Remote Sensing Data[§]

ABSTRACT

Quantification of forest parameters in different successional stages is important to assess the physical characteristics of the forest physiognomy. We modeled the forest attributes using both a parametric multiple linear regression analysis and neural networks approach, using Landsat ETM data acquired in 2000 (ETM00). We compiled sample plot data using forest inventory data collected from 1997 to 1998. There were 226 plots used to train the models and 112 plots were used for the validation. The remote sensing data (spectral values, vegetation indices, texture, etc) coupled with digital elevation model (DEM) were experimented and selectively used to model basal area, stem volume and above ground biomass (AGB). We investigated the possibility to estimate the forest attributes using Landsat ETM image from 2003 (ETM03) and applying multivariate alteration detection method. Our results showed that the mean texture index is strongly correlated with the forest attributes. We also show that neural networks perform better than multiple regression for predicting the forest attributes. The modeled basal area, stem volume and AGB varied from 10.7 – 15.1 m².ha⁻¹, 123.2 – 181.9 m³.ha⁻¹, and 132.7 – 185.3 Mg.ha⁻¹, respectively. The RMSE_r values of model fitting ranged from 11.2% to 13.3%, and the test dataset estimated slightly higher RMSE_r which varied from 12% to 14.1%. The transferability of the forest attribute models from the ETM00 data to ETM03 revealed favorable estimates, although the ETM03 shows considerably higher estimates than the ETM00.

Keywords: successional forests, forest attributes, tropical forests, Landsat ETM image, multiple linear regression, neural networks, basal area, stem volume, AGB

[§] Wijaya, A., V. Liesenberg and R. Gloaguen (2010). *Retrieval of Forest Attributes in complex Successional Forests of Central Indonesia: Modeling and Estimation of Bi-Temporal Data*. Forest Ecology and Management, Vol.259, Page 2315 – 2326.

6.1 Introduction

Indonesia has the third largest tropical forest after Brazil and the Democratic Republic of Congo around the globe (FAO, 2006). However, the rate of forest loss in Indonesia is accelerating rapidly. On average, about 1 million hectares per year of forests were cleared in the 1980s, rising to about 1.7 million hectares per year in the first part of the 1990s. Since 1996, deforestation appears to have increased to an average of 2 million hectares per year (FWI/GFW, 2002). Nonetheless, some efforts have been made to recover previously forested areas. Regeneration of vegetation is common in some areas following the completion of forest disturbances and/or deforestation. The intensity of regenerating processes applies differently within different forest environments depending on the climate, terrain conditions, soil fertility and nutrient contents, light intensities, characteristics of pioneer vegetation species, etc (Lu et al., 2003; Moran and Brondizio, 1998). The resulting landscape often consists of patches of successional forests and agricultural lands (Lu et al., 2003). This successional forest plays an important role in soil restoration through the accumulation of biomass, the buildup of litter and soil organic matter and other beneficial soil/plant interactions (Moran et al., 2000).

The successional forests might be distinguished based upon the time since the last abandonment, the average stand height and basal area, the stand volume and biomass density, and physiognomic characteristics (Moran and Brondizio, 1998; Uhl et al., 1988). The quantification of specific forest parameters in different successional stages is therefore important to assess the status of successional forest at different stages of their evolution. This information is also useful to better understand the relationship between successional stages and ecosystems change (Lu et al., 2003). Finally, the fundamental ecological properties impacting the functioning of many terrestrial processes, such as water and nutrient cycling and carbon sequestration need to be understood at different successional stages (Song et al., 2007).

Foody and Curran (1994) highlighted that basic information on the extent and the biophysical properties in different successional forest physiognomies are relatively poorly unknown. We propose to investigate such properties using multi-temporal remote sensing data.

This study focuses over a lowland tropical forest in East Kalimantan province, Indonesia. The island of Kalimantan includes the largest forest areas within this country (MoF, 2006). The Indonesian lowland tropical forest is among the world's richest in timber resources and biodiversity (FWI/GFW, 2002). Different forest physiognomies and successional stages,

such as tropical peat swamp forest, open forest, riparian forest, successional forests characterized with dense and very dense forest canopy, advanced secondary forest, and primary forest may be found in the Kalimantan forests (BFMP, 1997). Compared to the Brazilian Amazon (Li et al., 1994; Lucas et al., 2002; Lucas et al., 1993; Mausel et al., 1993; Moran and Brondizio, 1998; Moran et al., 1994; Steininger, 1996; Tucker et al., 1998), there are relatively few studies dealing with the characterization of successional forest stages in Kalimantan (Nykvist, 1996; Sist and Nguyen-Thé, 2002). Therefore, an attempt to assess the forest properties, which yields accurate information on the forest structure at different successional stages is of great interest. It is even highly appreciable, if the proposed method is reproducible, easy to implement and transferable to another forest region (Foody et al., 2003).

Remote sensing approaches become more popular as huge areas can be covered with less efforts and time, with regard to different sensor characteristics and limitations (Houghton et al., 2001; Lu, 2005; Lu, 2006). Different types of spaceborne remote sensing data employing optical sensor (Lu, 2005), synthetic aperture radar (SAR) data (Hajnsek et al., 2005) or light detection and ranging (LIDAR) data (Lefsky et al., 2002) have been explored to quantify forest biophysical properties. In the last decade, forest biomass assessment has become one of the major concerns for scientists working with remote sensing for forestry applications (Kuplich et al., 2005; Li et al., 2008; Lu, 2005; Lu, 2006; Luckman et al., 1996). Previous studies reported that the transferability of the models generated from remote sensing data is limited because these models are rather site specific (Foody et al., 2003; Houghton et al., 2001).

The application of coarse resolution data, e.g. Advanced Very High Resolution Radiometer (AVHRR) or Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, are feasible for regional mapping of forest biophysical properties (Hame et al., 1997), but the sensors with moderate resolution, such as Landsat or ASTER data, is required for more accurate estimation (Muukkonen and Heiskanen, 2005). Although the most accurate estimate is achieved using LIDAR data, the relatively small covered areas and relatively expensive acquisition are the major obstacle limiting the utility of this particular sensor in larger scale (Drake et al., 2003; Gillespie et al., 2004). Besides, the utility of SAR data for biophysical properties estimation is also widely studied (Cloude et al., 2008; Hajnsek et al., 2005; Isola and Cloude, 2001). Here, we used Landsat ETM data which is highly abundant and at the same time is underutilized, to develop the statistical models of forest attributes, namely basal area, stem volume and above ground biomass (AGB) in Indonesian forest environments.

A great number of studies has been undertaken to predict the forest structural properties using the Landsat ETM data (Lu and Batistella, 2005; Lu et al., 2004; Sader et al., 1989; Sales et al., 2007; Steininger, 2000; Steininger, 1996). These studies delivered mixed results in estimating the forest structure properties, especially those conducted in tropical forests (Foody et al., 2003; Ketterings et al., 2001; Rahman et al., 2005), where the vegetation structure is very complex and relatively few accurate field observation data are available. Thus, further study to model the forest attributes in this particular forest environment is still necessary. Any additional information might improve the correlations between the modeled forest attributes and remote sensing data (Foody et al., 2003; Kimes et al., 1998; Lu, 2005; Lu and Batistella, 2005). Therefore, in order to increase the estimation accuracy, we coupled Landsat ETM reflectance, with vegetation indices, image transform layers, texture features, and digital elevation model (DEM). Each feature was tested for its predictive potential and then selectively used to model forest attributes.

We modeled the forest attributes using (1) a parametric multiple linear regression analysis and (2) a neural networks approach. Both methods have been successfully used in the estimation of forest structure properties, using remote sensing (Ardo, 1992; Boyd et al., 2002; Hame et al., 1997; Hyypä et al., 2000). A possibility to estimate the forest attributes from bitemporal Landsat ETM data was also investigated by calibrating the radiometric properties of across date ETM image using an automatic relative radiometric calibration approach based on multivariate alteration detection (MAD) (Canty et al., 2004; Nielsen et al., 1998).

6.2 Data and Methods

6.2.1 Image acquisition and preprocessing

Satellite data used in this study includes Landsat 7 Enhanced Thematic Mapper plus (ETM+) and digital elevation model (DEM) from Shuttle Radar Topography Mission (SRTM) sensors. Two Landsat ETM data acquired on August 26, 2000 and May 31, 2003 were employed. The Landsat ETM data from 2000 (henceforth ETM00) was used to model the forest attributes, whereas the ETM of 2003 (ETM03) was used to investigate the estimations of forest attributes from across date ETM scene.

The ETM00 image was projected into WGS 84 datum and Zone 50N using Universal Transverse Mercator (UTM) projection. Preprocessing of the ETM00 data was conducted for correcting the atmospheric and topographic effects by minimizing the artifacts caused by the atmospheric attenuations, e.g. haze and irradiance scattering, and the terrain effects. Also, calculation of vegetation indices required the surface reflectance instead of DN values

or top of atmosphere reflectance, so that the image corrections were required. Atmospheric corrections were applied on the ETM00 using dark object subtraction (DOS) method proposed by Chavez (1988). A study conducted by Song et al. (2001) proposed different variations of DOS techniques. We experienced that the COST-DOS technique offered more preferable results with regard to the spectral responses of vegetated areas. Topographic corrections were implemented using C-Correction procedure assuming Lambertian effects on the earth surface (Riaño et al., 2003).

To update the existing GIS-based land cover map, we used the ETM00 data and additional ground truth data obtained from field work. A maximum likelihood supervised classification was used and the classification accuracy was assessed using an independent test dataset, resulted in 78% of overall accuracy and Kappa statistics of 0.75. We applied a majority analysis on the classification image, which is basically an attempt for removing minor spurious pixels within a large single class. Prior to the analysis, a 5×5 kernel size was selected, in which the center pixel was replaced with the dominant class label within the area covered by the moving window. This process was iterated for the entire image resulting in more homogenous classification results. The accuracy of the processed land cover map increased to 83% and Kappa of 0.80. Also, the classified image showed more favorable visual characteristics than did the unprocessed one. Ultimately, the 'updated' land cover map, consists of five land cover types, namely open forest (OF), riparian forest (RF), successional forest 1 (SF1), successional forest 2 (SF2) and advanced secondary forest (ASF), and were briefly explained in Tab. 6.1.

Tab. 6.1 Description of sample data showing number of plots, total stems, mean diameter at breast height (DBH), basal area and stem volume for different forest physiognomies

	Open forest (OF)	Riparian forest (RF)	Successional forest 1 (SF1)	Successional forest 2 (SF2)	Advanced secondary forest (ASF)
Number of Stems (ha^{-1})	221.7	255.3	250.7	266.1	244.7
DBH (cm)	29.1	31.4	35.5	35.0	43.8
Basal Area ($\text{m}^2.\text{ha}^{-1}$)	10.9	13.3	13.5	14.1	15.9
Stem Volume ($\text{m}^3.\text{ha}^{-1}$)	114.8	148.0	154.5	161.0	193.0
Biomass ($\text{Mg}.\text{ha}^{-1}$)	124.5	157.0	165.7	169.9	201.1
Number of plots (n = 338)	12	16	190	104	16
Canopy cover	< 50%,	< 50%,	60% - 75%	75% - 90%	over 90%
Descriptions	Mixture of pioneer species, low to medium tree size and shrubs, currently disturbed, exhibits noticeable marks of forest burning and clearing	Sparse forest dominated with slim and tall vegetation, located adjacent to the streams	Dense forest, logged over < 10 years, located in flat and moderate slope	Very dense forest, logged over between 10 - 20 years, located in moderate and highly steep regions	Advanced forest structure, closed canopy, logged over > 20 years, located mostly in highly steep region

To investigate the possibility for estimating the forest attributes from bitemporal ETM data and to test the robustness of the generated models, we used the ETM image from 2003. The ETM03 image was co-registered to the ETM00 data and resampled using nearest neighborhood algorithm. To normalize the radiometric properties of the ETM03, a relative radiometric normalization, the multivariate alteration detection (MAD) (Nielsen et al., 1998) was applied using the ETM00 as reference. The linear scale invariance of the MAD transformation was used to obtain invariant pixels for automatic relative radiometric normalization of the bitemporal Landsat ETM data (Canty et al., 2004). Invariant pixels were selected from the ETM00 and ETM03 subsets which show relatively no-change pixels. Linear combinations of the intensities were iterated to all six bands in both ETM subsets. These linear combinations in each paired band, or called MAD components (MAD_i) are invariant under linear transformations of the original image intensities. For radiometric normalization, we selected all pixel coordinates, which satisfy

$$\sum_{i=1}^n \left(\frac{MAD_i}{\sigma_{MAD_i}} \right)^2 < t, \quad \text{Eq. 6.1}$$

where σ_{MAD_i} is the standard deviation of each MAD components (MAD_i) and t is a decision threshold. Under the hypothesis of no-change pixels, the above sum of squares of standardized MAD variables is approximately chi-square distributed with N degrees of freedom. We therefore choose $t = \chi^2_{N, p} = 0.05$ where p is the probability of observing the value of t or lower. The selected pixels should correspond to truly invariant features so long as the overall radiometric differences between the two images can be attributed to linear effects. Since this method identified quite a large number of no-change pixels ($n = 105844$), half of selected pixels were randomly selected for performing the linear regression. The remaining pixels ($n = 52921$) were used for validating the regression model by performing paired t -tests for equal means and its counterpart F -tests for equal variances. Normalization on the basis of no-change pixels was performed to the entire ETM03 image by means of ordinary least square (OLS).

The SRTM DEM originally of 90 meter resolution was also registered with the ETM00 data and resampled using nearest neighborhood algorithm into 30 meter spatial resolution to allow the forest attributes modeling jointly with the ETM data. Slope angle and aspect were then computed from the resampled DEM.

6.2.2 Generation of ancillary remote sensing data

Additional remote sensing variables were calculated from the ETM00 surface reflectance. The application of simple ratios and both of traditional and complex vegetation indices is wide, ranging from soil moisture, vegetation monitoring to mineral deposits mapping (Jensen, 1996). This study, however, selected the ratios or indices which were proposed for the monitoring of green vegetation and the assessment of forest biophysical properties (Tab. 6.2). Most of these indices were calculated using a combination of red (ETM band 3), near infrared (ETM band 4), and middle infrared (ETM band 5) bands.

Besides, three bands of principal component analysis (PCA) and tasseled cap (TC) bands, i.e. brightness (TC1), greenness (TC2) and wetness (TC3) were calculated as well. We also include texture features based on the grey level co-occurrence matrix (GLCM). Prior to the computation of the texture features, a variance matrix calculated from each ETM band was analyzed, to find specific ETM band representing the highest variance of forested lands. We found the ETM band 5 (middle infra red band, $ETM_{mid} = 1650$ nm) revealed a substantial variance for different forest physiognomies. It means that this particular band may contain more information than other ETM bands, to develop the correlations between the ETM data and forest properties under study. The middle infra red band of ETM data was ultimately selected for the texture features generation using a 5×5 moving window. The texture layers were calculated in each direction with single shifting pixel and were quantified into a 64 gray levels. Eight texture features computed from mean (GLCM_MEAN), variance (GLCM_VAR), homogeneity (GLCM_HOMO), contrast (GLCM_CONT), dissimilarity (GLCM_DISS), entropy (GLCM_ENTR), second moment (GLCM_SECM), and correlation (GLCM_CORR) were produced.

Tab. 6.2 Various simple ratio, different NDVIs, complex vegetation indices and image transform layers computed in this study to generate the biomass and stem volume equations

Index under study	Formula	Description	Reference
Simple band ratios			
ETM 4/3	$ETM4/ETM3$	Measure of green vegetation and chlorophyll absorption bands (ETM _{mid} : 660nm; 830nm)	(Rouse et al., 1973)
ETM 5/3	$ETM5/ETM3$	Modified simple ratio (ETM _{mid} : 660nm; 1650nm)	(Lu et al., 2004)
ETM 5/4	$ETM5/ETM4$	Modified simple ratio (ETM _{mid} : 830nm; 1650nm)	(Lu et al., 2004)
ETM 5/7	$ETM5/ETM7$	Modified simple ratio (ETM _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)
ETM 7/3	$ETM7/ETM3$	Modified simple ratio (ETM _{mid} : 660nm; 2215nm)	(Lu et al., 2004)
Traditional vegetation indices			
NDVI	$(ETM\ 4 - ETM\ 3)/(ETM\ 4 + ETM\ 3)$	Measure of green vegetation cover (ETM _{mid} : 660nm; 830nm)	(Rouse et al., 1973)
ND53	$(ETM\ 5 - ETM\ 3)/(ETM\ 5 + ETM\ 3)$	Modified vegetation indices (ETM _{mid} : 660nm; 1650nm)	(Lu et al., 2004)
ND54	$(ETM\ 5 - ETM\ 4)/(ETM\ 5 + ETM\ 4)$	Modified vegetation indices (ETM _{mid} : 830nm; 1650nm)	(Lu et al., 2004)
ND57	$(ETM\ 5 - ETM\ 7)/(ETM\ 5 + ETM\ 7)$	Modified vegetation indices (ETM _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)

ND32	$(ETM\ 3 - ETM\ 2)/(ETM\ 3 + ETM\ 2)$	Modified vegetation indices (ETM+ _{mid} : (Lu et al., 2004) 560nm; 660nm)	
Complex vegetation indices			
ARVI	$(NIR + 2RED + BLUE)/(NIR + 2RED - BLUE)$	Enhancement of NDVI that is relatively resistant to atmospheric factors (Kaufman and Tanre, 1996) (ETM+ _{mid} : 485nm; 660nm; 830nm)	
EVI	$2.5 \times (NIR - RED)/(NIR - 6RED - 7.5BLUE + 1)$	Reduce the atmospheric influence and optimize the vegetation signal (Huete et al., 1997) (ETM+ _{mid} : 485nm; 660nm; 830nm)	
SAVI	$(NIR - RED) \times (1 + L)/(NIR + RED + L)$	Modified green vegetation index with an adjustment factor (Huete, 1988) (ETM+ _{mid} : 660nm; 830nm)	
MSAVI2	$\left((2NIR+1) - \sqrt{(2NIR+1)^2 - 8(NIR-2RED)} \right)/2$	Measure of vegetation that is less sensitive to atmosphere and soils (Qi et al., 1994) (ETM+ _{mid} : 660nm; 830nm)	
GEMI	$\varepsilon(1-0.25\varepsilon)-(RED-0.125)/(1-RED)$	Global environmental monitoring index (Pinty and Verstraete, 1991) that is insensitive to empirical atmosphere (ETM+ _{mid} : 660nm; 830nm)	
	Where		
	$\varepsilon=(2(NIR^2-RED^2)+1.5NIR+0.5RED)/(NIR+RED+0.5)$		
Image transform indices			
VIS123	$ETM1 + ETM2 + ETM3$	(ETM+ _{mid} : 485nm; 560nm; 660nm)	(Lu et al., 2004)
MID57	$ETM5 + ETM7$	(ETM+ _{mid} : 1650nm; 2215nm)	(Lu et al., 2004)
ALBEDO	$ETM1 + ETM2 + ETM3 + ETM4 + ETM5 + ETM7$	(ETM+ _{mid} : 485nm; 560nm; 660nm; 830nm; 1650nm; 2215nm)	(Lu et al., 2004)

6.2.3 Field observation data

Field observation data were compiled from the inventory data collected by the forest concession management from April 1997 to January 1998. We selected 338 sampling plot data (15m × 15m), and divided the data into training dataset (n = 226) and test data (n = 112), for modeling and validating the forest properties, respectively. The sample data, collected in 16 transects, were spanned along the NW – SE directions within the study area. In each measurement plot, diameter at breast height (DBH) ranging from 10 – 140 cm, number of stems, and geographical coordinates were recorded. In total 3078 trees were measured, using which tree basal area and stem volume were calculated. The estimation of tree volume considered the adjustment factor for different tree species. Above ground biomass (AGB) was subsequently estimated from tree diameter using an existing AGB allometric equation modeled for particular tropical lowland *dipterocarp* forests (Samalca, 2007).

$$AGB = \exp(-1.2495 + 2.3109 \times \ln(dbh)) \quad \text{Eq. 6.2}$$

The Eq. 6.2 was generated by destructively sampling 40 trees, to estimate dry weight of carbon contents. To date, field measurement based approach is considered as the most accurate method for forest biomass assessment (Lu, 2006).

All those parameter estimates were multiplied by the number of stems per hectare and for each sample plot the values were aggregated to calculate per hectare measures (i.e. basal area, stem volume and AGB per hectare) in each plot. The updated land cover map was subsequently used to characterize the land cover type of each sample plot, as this information was not collected by the forest concession management during the field campaign. Descriptions on the physical characteristics of the forest successional stages were briefly explained in Tab. 6.1.

Within the scope of this study, basal area, stem volume, and AGB from different successional forests were modeled. The basal area, defined as the cross section area of trees measured at breast height (~138 cm), is generally expressed as square units per unit area. Calculation of tree basal area is used as a measure of tree density and to determine percent stocking. Estimation of stem volume is important for mapping of standing stock and for forest inventory purpose, as it provides initial prediction on timber amount that could be commercially harvested. The AGB estimate, on the other hand, indicates the carbon accumulation in a forest region over the time, and information on total AGB estimated from each forest successional stage is useful to assess how different regeneration stages could have an effect on the forest as a source of carbon sink.

6.2.4 Statistical modeling

Estimation of the forest parameters from Landsat 7 ETM+ and SRTM DEM data employed a multi-linear regression analysis and neural networks. Variations of multispectral bands, vegetation indices, image transform bands, texture features, and DEM data were used as the predictors. The selection of the predictors was based on the correlations with the forest attributes under study. The regression analysis followed a forward stepwise regression procedure to estimate the model parameters,

$$y_j = ax_1 + bx_2 + \dots + gx_i + \varepsilon_i \quad \text{Eq. 6.3}$$

where the y_i is the forest parameters, the $a - g$ are the parameters of respective predictors (x_i), and ε_i is the error term. Several predictor variations were also tested to find the best fitted model for each forest attribute.

Neural networks were mainly used for the classification of remotely sensed image (Atkinson and Tatnall, 1997). This study demonstrates the utility of neural network for regression analysis to quantify the forest attributes. A possibility to combine various RS data (e.g. spectral data, vegetation indices, etc) considered in this study may certainly violate a normal-distribution assumption which is required by a linear regression method to predict the forest attributes accurately. Also, application of different RS data may compromise the underlying assumptions of a parametric approach which are required by the regression analysis (Boyd et al., 2002). Thus, a typical free assumptions approach, such as neural network, might provide better estimates than do conventional parametric approaches (Muukkonen and Heiskanen, 2005).

The neural networks were trained using Matlab version 2006a software (Demuth et al., 2006). This method works by aggregating the weighing of the input values and applying a threshold function into those values. The network training process is sometimes very time-consuming, and the selection of a faster and more efficient training algorithm is preferable. Instead of applying the back-propagation algorithm which is widely applied for remote sensing study but actually inefficient in training the networks (Atkinson and Tatnall, 1997), this study alternatively used the Levenberg-Marquardt algorithm in training the networks. The Levenberg-Marquardt algorithm is basically an iterative technique that locates the minimum of a function expressed as the sum of squares of nonlinear functions (Hagan and Menhaj, 1994). It is a standard technique for nonlinear least-squares problems combining the steepest descent and Gauss-Newton methods. When the current solution is far from the correct one, the algorithm behaves like a steepest descent method: slow, but guaranteed to

converge, and it becomes a Gauss-Newton method when the current solution is close to the correct one. The networks employed had one hidden layer with 14 hidden nodes between the input and output layers, i.e. between remote sensing data and forest properties. The hyperbolic tangent sigmoid and linear transfer functions were selected to estimate the hidden- and output neurons, respectively (Demuth et al., 2006). Before training the networks, the training dataset ($n = 226$) was randomly allocated into training (80%) and validation (20%) subsets. The training subset was used for training the neural networks, while the validation subset was used as a proxy of the training processes to avoid the over fitting problem.

6.2.5 Validation

The linear regression analysis and neural networks were applied to the test dataset ($n = 112$) to calculate the accuracy statistics. The accuracy statistics include the root mean square error (RMSE) and bias, and their relative counterparts $RMSE_r$ and $bias_r$ (Muukkonen and Heiskanen, 2005).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - j_i)^2} \quad \text{Eq. 6.4}$$

$$RMSE_r = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - j_i)^2}}{y_m} \times 100 \quad \text{Eq. 6.5}$$

$$Bias = \frac{1}{n} \sum_{i=1}^n (y_i - j_i) \quad \text{Eq. 6.6}$$

$$Bias_r = \left[\frac{1}{n} \sum_{i=1}^n (y_i - j_i) \right] y_m \times 100 \quad \text{Eq. 6.7}$$

Where j_i is the predicted value, y_i is the observed value, y_m is the mean of observed value and n is the number of plots in test dataset.

To examine if there were significant differences between the observed forest attributes and those estimated from the linear regression and neural networks, a non parametric Wilcoxon signed rank test was applied. We expect the results might reject the hypothesis that the mean values of the observed and modeled forest properties differ significantly.

6.3 Results

6.3.1 Distribution of forest attributes

The number of stems, basal area, stem volume and AGB ($n = 3078$) were plotted against DBH classes showing that 95% of sampled trees ($n = 2909$) have $DBH < 80$ cm, and more than 65% of the trees ($n = 2090$) have $DBH < 40$ cm (Fig. 6.1). The high abundance of small trees ($DBH < 40$ cm), however, contributes only to relatively low stem volume ($15.4 \text{ m}^3.\text{ha}^{-1}$) and AGB ($15 \text{ Mg}.\text{ha}^{-1}$), as compared to large trees ($DBH > 80$ cm) with mean stem volume and AGB of $72.6 \text{ m}^3.\text{ha}^{-1}$ and $72 \text{ Mg}.\text{ha}^{-1}$, respectively.

The small trees ($DBH < 40$ cm) were dominant in the OF and SS1 classes, and partly in the SS2 plots, which were mostly occupied by small and fast growing pioneer species. The pioneer trees and small light demanding species disappeared during the regenerating process due to natural thinning effects caused by species competition in pursuing limited nutrient contents and light intensities (Moran et al., 2000). The RF on the other hand, comprises of a mixture of shrubs, young trees, and tall and slim emergent palm trees in the top canopy layer which is similar with the OF and SS1. But the presence of the emergent palm trees in the canopy layer distinguishes the RF from the other forest classes.

Based on the tree sample data (not shown here), successional forests plots comprise of trees with mean diameter varies from 20 – 65 cm. The high variation in DBH explained the variations of stem volume ($60 - 222 \text{ m}^3.\text{ha}^{-1}$) and AGB ($61 - 234 \text{ Mg}.\text{ha}^{-1}$) in the successional forests. In the ASF plots, the mean tree diameter ranges from 27 – 77 cm. The stem volume and AGB in these particular plots varied from $142 - 214 \text{ m}^3.\text{ha}^{-1}$ and $168 - 227 \text{ Mg}.\text{ha}^{-1}$, respectively. These variations were common for natural forests especially those which are occupied by successional and secondary forests. Tree regenerating processes take place following the completion of forest harvesting, forest burning, and other forest disturbances. These processes which can continue for over 30 years are affected by various dependent and independent aspects, e.g. anthropogenic factors, drought, disease, etc (Bischoff et al., 2005; Vieira et al., 2003).

The stem volume and AGB exponentially increased with the DBH (Fig. 6.1). These were because the biomass has an exponential relationship with the tree diameter (DBH), and the stem volume is basically a square function of DBH. Keeping all other conditions constant, the stem volume and AGB presented a linear relationship with each other.

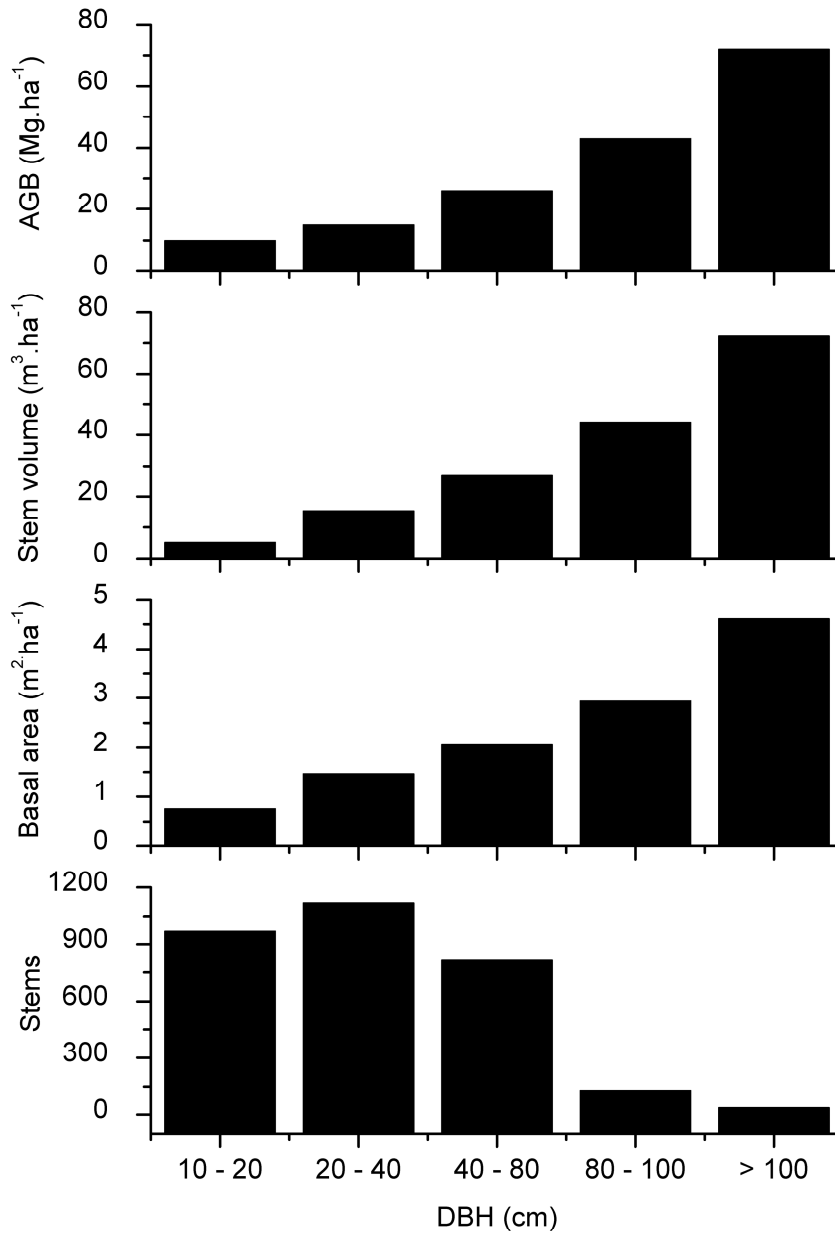


Fig. 6.1 Distribution of forest properties in different diameter at breast height (DBH)

6.3.2 Correlations

Pearson correlation analyses showed negative correlations between the forest attributes and remote sensing data (Tab. 6.3). The GLCM mean texture feature had relatively strong correlation with basal area ($r = -0.517, p < 0.01$), stem volume ($r = -0.669, p < 0.01$) and AGB ($r = -0.544, p < 0.01$), whereas elevation of DEM showed the highest correlation with stem number ($r = -0.217, p < 0.01$). An attempt to increase the correlations were experimented by generating the GLCM texture features using larger window size, e.g. 11×11, but we found low correlations with the forest properties data. The reflectance of ETM band 4 (NIR), band 5 (MIR) and band 7 showed relatively moderate correlations with all forest attributes

except stem number (correlation coefficient r ranged from -0.298 to -0.418). For simple band ratios, only SR53 and SR57 significantly correlated with the forest attributes at $p = 0.01$, although the correlations were relatively weak ($-0.195 < r < -0.147$). The traditional vegetation indices showed almost no correlations with the forest properties, except for ND53 that significantly correlated with basal area ($r = -0.147, p < 0.01$). From the complex vegetation indices, only GEMI showed significant correlations with all forest attributes except stem number ($-0.368 < r < -0.293$). Also, the image transform indices (i.e. VIS123, MID57 and Albedo) showed relatively good significant correlations with the forest attributes except number of stem ($-0.443 < r < -0.201$). The two first principal components were significantly correlated with basal area, stem volume and AGB ($-0.426 < r < -0.323$). Similarly, the tasseled cap (TC) bands also showed significant correlations with forest properties ($-0.427 < r < -0.229$), whereas only GLCM mean texture showed significant correlations with the forest attributes ($-0.669 < r < -0.163$). In general, most of the RS data had relatively low correlations with the forest attributes ($r < 0.4$), except for ETM band 5, MID57, Albedo, PC1, TC1 and GLCM mean, which showed relatively strong correlations with the forest properties ($-0.7 < r < -0.4$).

6.3.3 Statistical models

Application of a stepwise multiple regression analysis suggested ETM B4, ETM B5, ETM B7, DEM elevation, PC1, TC3 and GLCM mean texture as the predictors in the forest attributes modeling. This combination of variables significantly predicted basal area ($F = 15.7, p < .001$), stem volume ($F = 26.9, p < .001$), and AGB ($F = 15.7, p < .001$). The beta weights (β), presented in Tab. 6.3, suggest that the ETM B5 and TC3 contribute most for predicting the basal area and AGB, whereas the ETM B4 and DEM elevation have more contribution for the stem volume estimation. The explanatory power r^2 values of the basal area, AGB, and stem volume were 0.34, 0.34 and 0.46, respectively. These indicate that 34% of the variances in basal area and AGB were explained by the models. Similarly, 46% of the stem volume variance was explained by the regression model. The $RMSE_r$ values varied between 11.5% and 13.3%, assessed from the training dataset, and between 12.2% and 14.4% when the models were applied to the test dataset. In general, for all forest attributes the $RMSE_r$ values of the testing data were higher than those of the training data.

Tab. 6.3 Correlations between remote sensing data and forest biophysical parameters

	Stem number (stems ha ⁻¹)	Basal area (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	AGB (Mg ha ⁻¹)		Stem number (stems ha ⁻¹)	Basal area (m ² ha ⁻¹)	Volume (m ³ ha ⁻¹)	AGB (Mg ha ⁻¹)
<i>Landsat ETM Reflectance</i>					<i>Complex vegetation indices</i>				
B1	-.092	-.197(**)	-.250(**)	-.246(**)	SAVI	-.070	-.112(*)	-.101	-.076
B2	-.092	-.220(**)	-.278(**)	-.275(**)	MSAVI2	-.060	-.088	-.075	-.052
B3	.000	-.122(*)	-.211(**)	-.194(**)	GEMI	.108(*)	.293(**)	.368(**)	.313(**)
B4	-.092	-.298(**)	-.395(**)	-.336(**)	<i>Image transform indices</i>				
B5	-.132(*)	-.337(**)	-.418(**)	-.375(**)	VIS123	-.066	-.201(**)	-.276(**)	-.267(**)
B7	-.142(**)	-.304(**)	-.390(**)	-.349(**)	MID57	-.142(**)	-.339(**)	-.425(**)	-.382(**)
<i>Digital elevation model (DEM)</i>					Albedo	-.121(*)	-.339(**)	-.443(**)	-.394(**)
Elevation	-.217(**)	-.144(**)	-.009	-.160(**)	<i>Principal component analysis (PCA)</i>				
Slope	.021	.074	.082	.082	PC1	-.112(*)	-.323(**)	-.426(**)	-.383(**)
<i>Simple band ratios</i>					PC2	.102	.307(**)	.399(**)	.340(**)
SR	-.092	-.151(**)	-.137(*)	-.109(*)	PC3	-.075	-.083	-.077	-.085
SR53	-.123(*)	-.195(**)	-.177(**)	-.160(**)	<i>Tasseled cap (TC) transformation</i>				
SR54	-.064	-.089	-.082	-.093	TC1	-.107(*)	-.323(**)	-.427(**)	-.374(**)
SR57	.026	-.043	-.037	-.037	TC2	-.064	-.229(**)	-.300(**)	-.241(**)
SR73	-.129(*)	-.162(**)	-.147(**)	-.130(*)	TC3	.133(*)	.309(**)	.380(**)	.338(**)
<i>Traditional vegetation indices</i>					<i>GLCM texture features</i>				
NDVI	-.068	-.101	-.085	-.063	GLCM_MEAN	-.163(**)	-.517(**)	-.669(**)	-.544(**)
ND53	-.099	-.147(**)	-.129(*)	-.111(*)	GLCM_VAR	.032	-.068	-.067	-.035
ND54	-.065	-.091	-.084	-.095	GLCM_HOMO	.074	.112(*)	.081	.078
ND57	.029	-.041	-.034	-.032	GLCM_CONT	.021	-.089	-.100	-.063
ND32	.108(*)	.108(*)	.061	.083	GLCM DISS	-.057	-.126(*)	-.099	-.085
<i>Complex vegetation indices</i>					GLCM_ENTR	-.083	-.066	-.011	-.016
ARVI	-.062	-.107(*)	-.099	-.073	GLCM_SECM	.099	.076	.011	.027
EVI	-.090	-.048	.014	-.002	GLCM_CORR	.018	.014	-.020	-.016

Tab. 6.4 Statistics and validation of linear regression analysis

Dependent variable	Predictor	Unstandardized coefficients		Standardized coefficients (β)	Train dataset				Test dataset		
		B	Standard error		F	R ²	RMSE	RMSEr (%)	RMSE	RMSEr (%)	Bias
Basal area	(Constant)	66.366	17.169								
	B4	-.238	.145	-.285							
	B5	-4.420	2.195	-2.706							
	B7	-3.342	1.680	-1.297	15.668 (p<0.001)	0.34	1.83	13.31	1.96	14.37	-0.135
	Elevation	-.009	.002	-.228							
	PC1	187.235	89.988	1.435							
	TC3	-518.050	261.934	-2.890							
	GLCM mean	-1.057	.160	-.475							
Stem volume	(Constant)	725.369	169.980								
	B4	-43.512	21.728	-2.385							
	B5	-34.542	16.632	-1.201							
	B7	1724.653	890.919	1.184	26.997 (p<0.001)	0.46	18.09	11.51	18.91	12.17	-2.781
	Elevation	-5281.835	2593.258	-2.639							
	PC1	-2.404	1.440	-.258							
	TC3	-.043	.024	-.097							
	GLCM mean	-14.842	1.581	-.597							
AGB	(Constant)	585.102	207.272								
	B4	-.918	1.756	-.090							
	B5	-29.092	26.495	-1.458							
	B7	-21.991	20.281	-.699	15.668 (p<0.001)	0.34	22.06	13.16	21.93	13.28	-3.100
	Elevation	-.113	.029	-.231							
	PC1	1051.897	1086.380	.660							
	TC3	-3339.196	3162.200	-1.525							
	GLCM mean	-12.586	1.927	-.463							

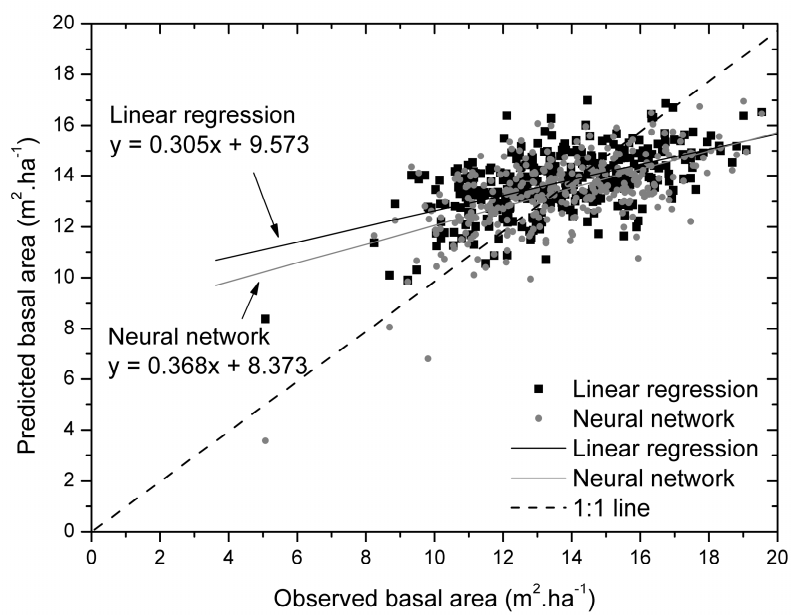
The explanatory power (r^2) and error estimates of neural networks, presented in Tab. 6.4, show that the neural networks models explain 36% of basal area, 50% of stem volume and 38% of AGB variances. Similar with the multiple regression analysis, the neural networks estimated higher $RMSE_r$ for the testing data than for the training data.

Comparison between the regression analysis and neural networks (Tab. 6.4 and Tab. 6.5) found that the models developed with a non parametric neural networks (NN) approach performed slightly better than the multiple regression (MLR) method for predicting the basal area ($r^2_{(NN)}=0.36$ vs $r^2_{(MLR)}=0.34$, $RMSE_{r(NN)}=14.14$ vs $RMSE_{r(MLR)}=14.37$) and stem volume ($r^2_{(NN)}=0.50$ vs $r^2_{(MLR)}=0.46$, $RMSE_{r(NN)}=11.99$ vs $RMSE_{r(MLR)}=12.17$). The neural networks model, however, resulted in better explanatory power ($r^2_{(NN)}=0.38$ vs $r^2_{(MLR)}=0.34$) but slightly higher error estimate than the multiple regression model ($RMSE_{r(NN)}=13.76$ vs $RMSE_{r(MLR)}=13.28$), for predicting the AGB. This was likely due to the over fitting problem during the networks training processes. The biases estimated from the neural networks were lower than those of the regression models with an exception of the basal area.

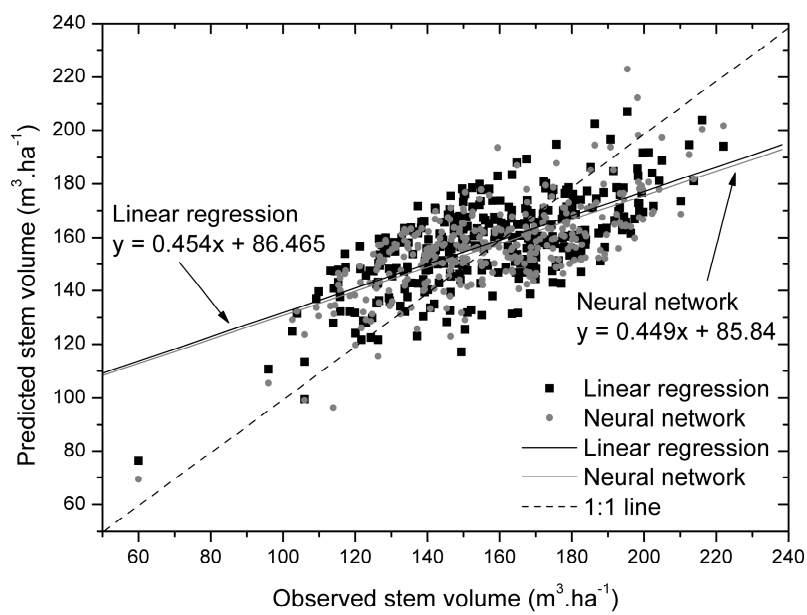
Tab. 6.5 Statistics and validation of neural networks regression analysis

Dependent variable	Train dataset			Test dataset		
	R^2	RMSE	$RMSE_r$ (%)	RMSE	$RMSE_r$ (%)	Bias
Basal area	0.36	1.81	13.17	1.93	14.14	0.225
Stem volume	0.50	17.58	11.18	18.63	11.99	-1.321
AGB	0.38	21.68	12.94	22.72	13.76	1.183

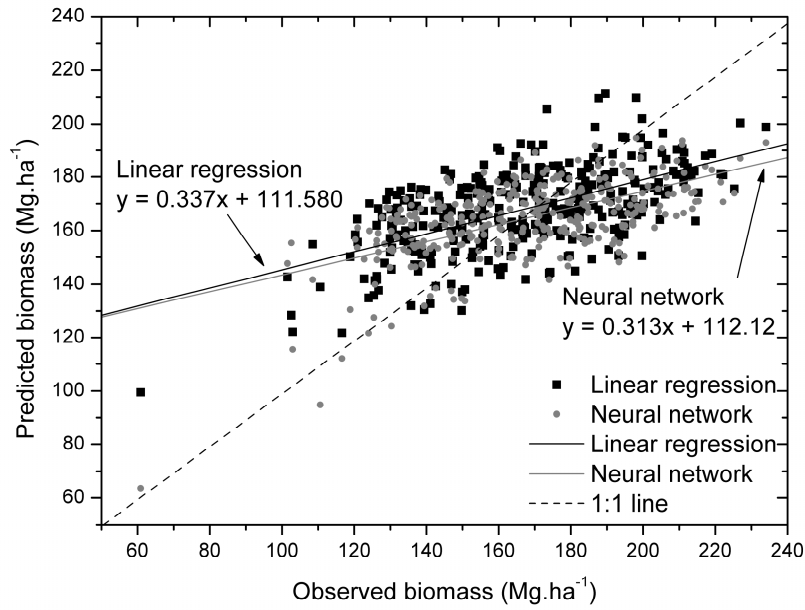
The predictions of the neural networks and the regression models are quite close to each other (Fig. 6.2). The models, however, tend to overestimate lower forest attributes values and vice versa. The reversal estimation points for the basal area, stem volume and AGB are $\sim 14 \text{ m}^2.\text{ha}^{-1}$, $\sim 160 \text{ m}^3.\text{ha}^{-1}$, and $170 \text{ Mg}.\text{ha}^{-1}$, respectively. These points indicate the asymptotic problem of the Landsat ETM data, which are not sensitive to model higher forest properties values. Some studies also reported similar problem, especially when Landsat data were used for estimating forest biomass (Lu, 2005; Steininger, 2000).



(a)



(b)



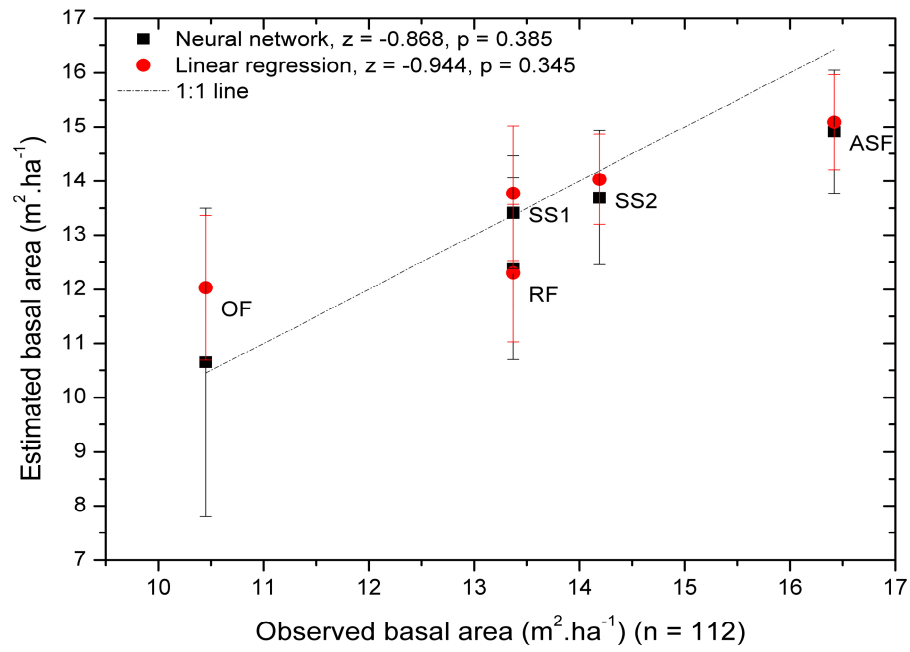
(c)

Fig. 6.2 The linear regression and neural network estimations of basal area (a), stem volume (b) and biomass (c) against the observed data. The forest attributes include complete dataset ($n = 338$)

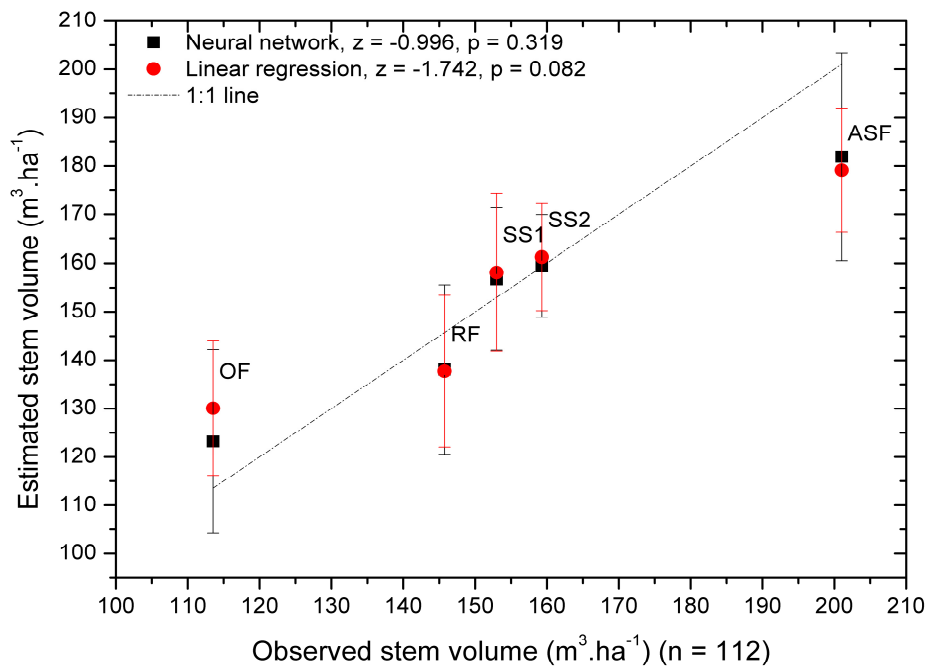
6.3.4 Results evaluation

The non-parametric Friedman test showed that at the p -value = 0.05, the basal area ($X^2 = 7.1$ with p -value 0.025), stem volume ($X^2 = 10.7$ with p -value 0.005) and AGB ($X^2 = 23.5$ with p -value < 0.001) of the multiple linear regression estimates, neural networks, and the observed data were significantly different.

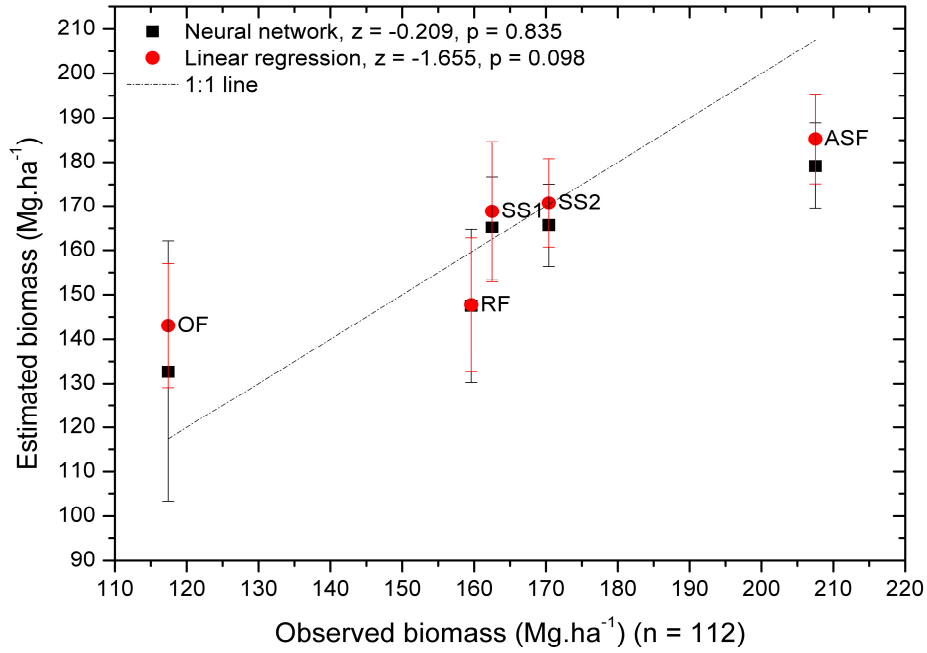
The forest attribute estimates were compared by forest classes with the observed data. The Wilcoxon signed rank sum test presented in Fig. 6.3 suggests that at the p -value of 0.01 level of significance, there were no significant differences between the multiple linear regression and observed data (p -value from 0.082 to 0.345), and between the neural networks and observed data (p -value from 0.319 to 0.835). The multiple linear regression and neural networks were also statistically compared, and the results highlighted significant differences in the basal area (z -test = -3.9, p -value < 0.001) and AGB (z -test = -4.9, p -value < 0.001), but not the stem volume (z -test = -2.3, p -value = 0.019).



(a)



(b)



(c)

Fig. 6.3 Comparison of multiple linear regression and neural networks predictions of basal area (a), stem volume (b) and biomass (c) against observed data by forest classes and z-test statistics of Wilcoxon signed rank sum test for paired dependent sample data. The regression and neural networks were tested separately against the observed test dataset (n = 112)

In general, the statistical test results suggest that the distributions of the forest attributes from the regression analysis, neural networks and observed data were significantly different. This difference is due to the statistical differences between the regression analysis and neural networks. Also, there was evidence that the neural networks performed significantly better than the regression analysis for predicting the basal area and AGB. For the stem volume, the neural networks were not significantly better than the regression analysis.

6.3.5 Forest attributes prediction from Bitemporal ETM data

To investigate the possibility for estimating the forest attributes from bitemporal ETM image, the ETM03 data was radiometrically normalized using the ETM00 as the reference. The ordinary least square regression analysis was employed to model the randomly selected invariant pixels from both ETM images (n = 105844). The correlations between both ETM bands, presented in Tab. 6.6, ranged from 0.68 (ETM band 1) to 0.89 (ETM band 5). Also, the RMSE calculated from independent test dataset (n = 52921) varied from 0.16 (ETM band 1) to 1.19 (ETM band 4).

Tab. 6.6 Ordinary least square regression on randomly selected training pixels ($n = 105844$) and the root mean square error (RMSE) calculated from the testing pixels ($n = 52921$); α is the fitted intercept, β is the fitted slope, σ_α and σ_β are the standard deviation, and r is the correlation coefficient.

ETM Band	α	σ_α	β	σ_β	r	RMSE
1	-3.71	0.02	0.10	0.0003	0.68	0.16
2	-8.16	0.03	0.27	0.0006	0.80	0.32
3	-2.86	0.02	0.19	0.0006	0.70	0.29
4	-3.10	0.04	0.32	0.0005	0.89	1.19
5	-1.18	0.02	0.20	0.0003	0.89	0.65
7	-0.48	0.01	0.25	0.0006	0.81	0.45

Tab. 6.7 Comparison of mean intensities of randomly selected testing pixels ($n = 52921$) for the ETM 2003 before and after normalization against the image reference (ETM 2000), with paired t -tests for equal means (p -value = 0.01). The statistical tests were calculated from the uncalibrated ETM03 and ETM00 data

ETM Band	1	2	3	4	5	7
ETM03 mean (uncorrected)	63.98	46.09	33.85	73.09	53.35	23.95
ETM03 mean (normalized)	2.67	4.40	3.41	20.55	9.52	5.61
ETM00 mean (reference)	2.67	4.39	3.41	20.54	9.52	5.61
t	1.75	-1.21	1.08	-1.75	0.69	-0.03
p	0.08	0.23	0.28	0.08	0.49	1.00

Tab. 6.8 Comparison of variances of the testing pixels ($n = 52921$) for the ETM 2003 before and after normalization against the image reference (ETM 2000), with F -tests for equal variances (p -value = 0.01). The statistical tests were calculated from the uncalibrated ETM03 and ETM00 data

ETM Band	1	2	3	4	5	7
ETM03 variance (uncorrected)	2.31	2.78	2.57	61.94	39.96	6.63
ETM03 variance (normalized)	0.02	0.21	0.09	6.48	1.61	0.43
ETM00 variance (reference)	0.05	0.30	0.17	7.74	2.01	0.62
F	2.14	1.45	1.93	1.19	1.25	1.44
p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Comparison of mean intensities presented in Tab. 6.7 showed that the mean differences between the normalized ETM03 and ETM00 bands were between 0.0 and 0.1. The t-test for equal means revealed that at the p-value of 0.01 level of significance, the means of normalized ETM03 and the reference image, ETM00, were not significantly different (p-value from 0.08 to 1.0). Upon band-wise variances comparison using the F-tests statistics (at the p-value of 0.01 level of significance) however, there was no sufficient evidence that the variances were equal (p-value < 0.001) (Tab. 6.8).

The forest attributes predictions showed that the calibrated ETM03 data estimated 64%, 63% and 44% of higher basal area, stem volume and AGB, respectively, than predicted from the ETM00 (Tab. 6.9). The ETM00 found $14 \pm 1.2 \text{ m}^2.\text{ha}^{-1}$ of the basal area, distinctly lower than the ETM03 prediction ($23 \pm 1.7 \text{ m}^2.\text{ha}^{-1}$). Also, the stem volume ($258 \pm 20.3 \text{ m}^3.\text{ha}^{-1}$) and AGB ($242 \pm 18.6 \text{ Mg}.\text{ha}^{-1}$) predicted from the ETM03 were markedly higher than those predicted from the ETM00 data (stem volume $158 \pm 16.8 \text{ m}^3.\text{ha}^{-1}$ and AGB $168 \pm 15.7 \text{ Mg}.\text{ha}^{-1}$). For all forest attributes, the coefficients of variations ranged from 7% to 11%.

Tab. 6.9 Forest attributes estimated from the sampling plots locations using the predictors of ETM00 and calibrated ETM03. The forest attributes were predicted using the regression models generated using the ETM00 data.

	Basal area ($\text{m}^2.\text{ha}^{-1}$)		Stem volume ($\text{m}^3.\text{ha}^{-1}$)		AGB ($\text{Mg}.\text{ha}^{-1}$)	
	ETM00	ETM03	ETM00	ETM03	ETM00	ETM03
Mean	14	23	158	258	168	242
Min	8	10	76	51	100	52
Max	17	30	207	299	211	304
Range	9	20	131	248	111	252
SD	1.2	1.7	16.8	20.3	15.7	18.6
CV	9%	7%	11%	8%	9%	8%

To better understand the underlying factors that caused remarkable increase of the ETM03 forest attributes, land cover classes and above ground biomass (AGB) estimations from both ETM00 and calibrated ETM03 subsets are displayed (Fig. 6.4). Prior to the comparison, non forest classes were masked out, so that only AGB on the forest classes under study were considered. The ETM00 found relatively lower AGB in the RF and SS1 classes, and moderate AGB level in the areas close to the logging roads, which were mainly classified as the SS2 forest (Fig. 6.4a and Fig. 6.4b). Similarly, the ETM03 data showed the same trends, except that the AGB estimates were higher than those predicted from the ETM00 (Fig. 6.4b, Fig. 6.4d). Also, both ETM images predicted a great number of forest biomasses in the ASF class.

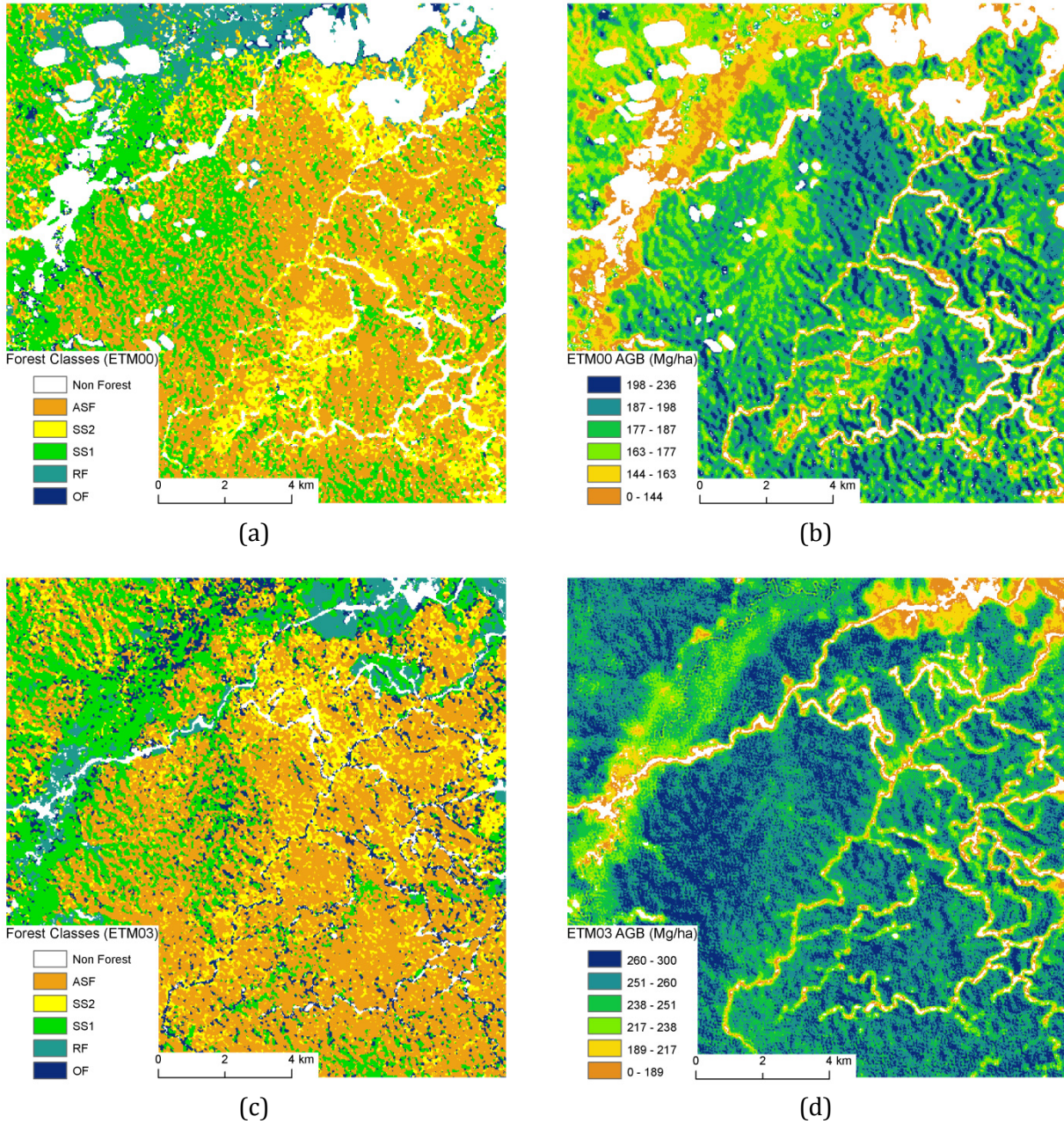


Fig. 6.4 Comparison of bitemporal ETM data showing the subsets of ETM00 land cover classification (a) with the ETM00 above ground biomass (AGB) estimates (b), and the ETM03 land cover map (c) and the AGB predictions of the radiometrically calibrated ETM03 image (d)

Comparison of the ETM00 and ETM03 land cover maps found more ASF pixels in the ETM03 than the ETM00, which were previously classified as the SS2 class by the ETM00 (Fig. 6.4a, Fig. 6.4c). This implies constant forest regenerating processes within this period. Generally, both ETM images were in agreement, explaining that the AGB increased with the forest regenerating processes, from the least to the more complex vegetation structures (i.e. from open forest (OF) to advanced secondary forest (ASF) class).

6.3.6 Discussion

6.3.6.1 Correlations

Texture features generated from the GLCM mean texture lead to the strong coefficient correlations with the forest attributes, except for the stem number. The correlation magnitudes were greater than those of the ETM multispectral bands, which normally were used to characterize the attributes of green vegetation. To facilitate better understanding about the relation of the mean texture and the forest attributes by each forest class, we plotted the mean texture feature against the observed AGB (Fig. 6.5). The ASF pixels were clustered on lower mean texture, whereas higher texture values mostly belong to the OF pixels. The SS1 and SS2 pixels are highly mixed and hardly to discriminate from this individual texture feature. Thus, the inclusion of other input variables, such as ETM spectral, vegetation indices, etc, is useful for generating more robust equations of the forest attributes, as performed in this study.

The strong correlation between the mean texture and the AGB as well as other forest attributes is probably due to the smoothing effects generated from the texture features that include the neighboring pixels in its calculation. Results of the present work are similar with a previous study that found the utility of texture features is useful for removing the shadow effects on broadleaf and/or large trees (Lu, 2005). That study highlighted a strong correlation between the entropy texture and AGB in uneven tropical forests of the Amazon. Another study conducted by Kuplich et.al. (2005) suggested the contrast texture to improve the correlation with the forest biomass. Negative correlations between the forest attributes and the ETM multispectral bands, particularly near infra red band were also confirmed by previous study (Lu et al., 2004). Another study by Lu and Batistella (2005) also found negative correlation between the texture features and the AGB, which was confirmed in this study.

6.3.6.2 Assessment of forest attributes estimation

Remote sensing based estimates have potential to predict the dynamics of forest attributes over large forest region with less efforts, time and cost (Lu, 2006). However, the accuracy of the estimates is somehow questionable, as it considerably depends on the quality of remote sensing data and its relationship with forest attributes being modeled (Foody et al., 2003). For instance, attempts to estimate above ground biomass using the Landsat data found high uncertainties which were around 30 – 40% (Lu et al., 2004; Lu, 2006; Sales et al., 2007). This study, however, found the RMSE of the AGB around 13% - 14% lower than the previous study. Given the fact that at some extent the AGB also correlates with the basal area and

stem volume, we can assume that the error estimates of these forest properties are also lower than in those studies. The relatively low error estimate can be explained from different perspectives, like the satellite data quality, selection of the models' predictors, the fitness of selected regression models (i.e. multiple regression and neural networks), or combination of those factors.

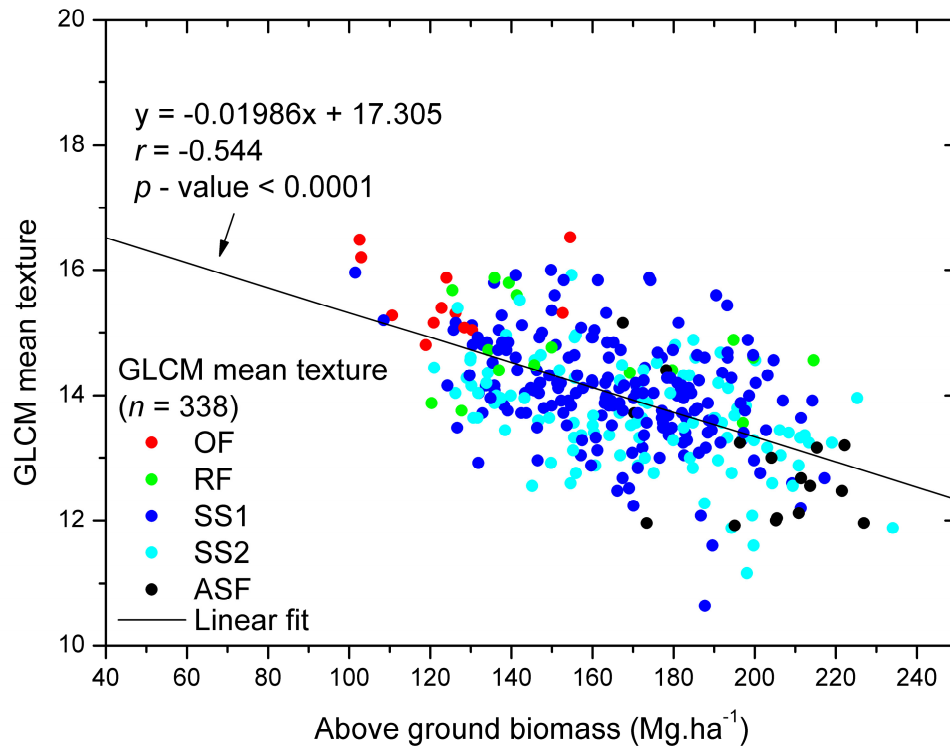


Fig. 6.5 Distribution of GLCM Mean Texture of Different Land Cover Type.

As explained earlier, estimation of the forest attributes employing Landsat ETM data was saturated at higher forest attributes and lowered the coefficient correlations. Beyond the saturation point, the ETM data is no longer sensitive to the forest attribute patterns. Several underlying factors might cause this problem, namely the size of sampling plot that was not designed to be related with spaceborne data, or the saturation from dense leaf canopies that restricts the forest attributes into a low level when the Landsat ETM was used (Anaya et al., 2009). Better results may be achieved either by applying laser scanner (LIDAR) data, SAR radar images, or high spatial resolution satellite imagery, e.g. IKONOS and Quickbird, which is beyond the scope of this study.

The correlation magnitudes between the ETM data and the forest attributes were relatively weak, although statistically the correlations were significant. For this particular case, the modelling of ETM data as a basis for the forest attributes prediction is problematic, as it may

affect the robustness of the developed models. Application of the non parametric method based on neural networks was used to evaluate the suitability of the multiple linear regression approach in estimating the forest attributes under these circumstances. We found the predictive ability of both methods seems comparable, although the neural network was slightly better than the regression analysis (Tab. 6.4 and Tab. 6.5). Also, the robustness of the regression models was tested on across date ETM image for predicting the AGB. The results obtained were markedly reasonable (Fig. 6.4).

Comparison of both land cover maps (i.e. ETM00 and ETM03) highlighted remarkable regenerating processes on the forested lands, which was explained by the presence of a great number of the ASF pixels in the ETM03 map. Atmospheric effects remaining in the ETM03 that were not compensated by the MAD normalization might explain some of the AGB differences from both ETM dates. Besides, severe El Nino drought and forest fires in 1997/1998 caused huge burned forest areas in 2000 and suppress the AGB into the lowest level as estimated from the ETM00. The increase of AGB in 2003 (i.e. the ETM03 estimate) was mainly due to forest regenerating processes following the completion of the fire event.

In spite of the results obtained in this study, further study may be conducted under other environmental conditions in order to test the models on different satellite sensors, such as MODIS, for regional mapping of forest attributes. It is a possible alternative to assess the forest attributes at larger scale given limited field data.

6.4 Conclusions

Generation of the GLCM texture features lead to relatively strong correlations between the mean texture with the basal area ($r = -0.517$), stem volume ($r = -0.669$) and above ground biomass (AGB) ($r = -0.544$). The correlation magnitudes were greater than the ETM reflectance ($r \leq -0.418$), simple ratio and vegetation indices ($r \leq -0.368$), image transform layers ($r \leq -0.443$), PCA ($r \leq -0.426$) and TC bands ($r \leq -0.427$). The elevation DEM showed the highest correlation with the number of stem ($r = -0.217$).

The forest attributes, modeled from the multiple linear regression (MLR), neural networks (NN) and the observed data were jointly tested using the Friedman statistical test, and the result showed that the data distribution were significantly different (p -value > 0.05). The Wilcoxon signed ranked test proved that this difference was due to statistical differences between the MLR and the NN estimates. The neural networks performed significantly better than the regression model for predicting the basal area and AGB, but not for the stem volume that showed no significant differences between both methods. The forest attributes increased with the advancement of canopy layers and vegetation complexity, from the least

to the most complex structure (i.e. the open forest (OF) class to the advanced secondary forest (ASF) class). The modeled basal area, stem volume and AGB varied from 10.7 – 15.1 m².ha⁻¹, 123.2 – 181.9 m³.ha⁻¹, and 132.7 – 185.3 Mg.ha⁻¹, respectively, depending on the forest types identified over the study area. The RMSE_r values of model fitting ranged from 11.2% to 13.3%, whereas the test dataset estimated slightly higher RMSE_r, which varied from 12% to 14.1%. These error estimates were relatively lower as compared to previous studies (Lu et al., 2004; Lu, 2006; Sales et al., 2007).

Normalization of the ETM03 data with the MAD method found similar mean intensities between the normalized image and the reference ETM00 image, but the variances were statistically different. The AGB were modeled using the calibrated ETM03 data as predictors, and the estimates were considerably higher than those predicted from the ETM00. Comparison of ETM00 and ETM03 land cover maps highlighted remarkable regenerating processes on the forested lands, which was explained by the presence of a great number of the ASF pixels in the ETM03 land cover map.

Chapter 7

Retrieval of Forest Biomass Using Mosaic SAR data^{**}

ABSTRACT

This study demonstrates the utility of mosaic ALOS PALSAR data for the estimation of above ground biomass and stem volume in tropical lowland forest of Kalimantan, Central Indonesia. The horizontal - horizontal (HH) polarization pair, horizontal - vertical (HV) pair and ratio of HV/HH polarizations were correlated with the forest properties using polynomial empirical models. The HV and HV/HH polarizations were best fitted for estimating the AGB and stem volume, respectively. The results showed that mosaic ALOS PALSAR data may be used as initial predictions of the AGB and stem volume, and the proposed approach have potential to apply for global datasets.

Keywords: ALOS PALSAR, radar backscatter, above ground biomass, stems volume

^{**} This chapter is based on:

Wijaya, A. and R. Gloaguen. (2009). *Fusion of ALOS Palsar and Landsat ETM Data for Land cover Classification and Biomass Modeling using Non-Linear Methods*. In Proceeding of International Geosciences and Remote Sensing Symposium (IGARSS), 13 – 17 July 2009, Cape Town, South Africa; and

Wijaya, A. (2009). *Evaluation of Mosaic ALOS Palsar Data for Estimating Above Ground Biomass and Stem Volume: A Case Study of Tropical Lowland Forest of Indonesia*. Jurnal Geografi, University of Indonesia, January 2009.

7.1 Introduction

Above ground biomass (AGB) and stem volume are two parameters which may be used to characterize current state of a forest. The estimation of these parameters is useful for environmental purpose as well as for forest inventory. The use of SAR data for estimating stem volume and AGB have been common practice for remote sensing study (Fransson and Israelsson, 1999; Luckman et al., 1996), as a previous study showed that optical data (e.g. Landsat ETM) has lack of correlations with the AGB and stem volume due to the sensor problem to penetrate through forest canopy cover (Lu et al., 2004). Alternatively, correlation with vegetation indices generated from those optical remote sensing data were usually used to assess the forest biophysical properties applying linear regression or neural network methods, but the results are questionable due to generalization and transferring ability problems (Foody et al., 2003).

SAR data is proved effectively for estimating above ground biomass and stem volume using an empirical model (Fransson and Israelsson, 1999; Luckman et al., 1996). Similar with the optical data, SAR data are also saturated at higher stem volume and AGB, but the capability to penetrate to cloud cover is the main advantage of this data compared to the optical data, particularly when the study focuses in tropical regions. This study experimented with a 50 meter spatial resolution of ALOS PALSAR mosaic data to assess the stem volume and AGB in tropical forest of Indonesia. The mosaic PALSAR data is freely available at the Kyoto and Carbon Initiatives Project website.

The present study contributes to develop empirical models for estimating the AGB and stem volume using radar backscatter of mosaic ALOS PALSAR data as the predictors. Relationship between HH, HV, HH/HV and HV/HH polarizations with the AGB and stem volume was analyzed. Allometric equation models for both forest parameters were estimated afterwards. The suitability of ALOS PALSAR data for estimating the AGB and stem volume is discussed.

7.2 Data and Methods

We used two sample datasets collected in 1997-1998 and 2003. The first dataset consists of 1512 temporary sample plots, and was used to generate the stem volume model using basal area (BA) as the predictor. The tree basal area was estimated and up scaled into per hectare measure. In total, 13050 trees were measured in the 1997-1998's dataset.

The 2003 dataset, collecting 38 circular sample plots was used to estimate stem volume based on the allometric equation developed from the 1997-1998 dataset. Above ground

biomass was estimated applying diameter at breast height (dbh) – dry weight AGB equation (Eq. 7.1) proposed by Samalca (2007).

$$AGB = -1.2495 + 2.3109 \times \ln(dbh) \quad \text{Eq. 7.1}$$

The 50-m resolution orthorectified mosaic ALOS PALSAR data was used to estimate the AGB and stem volume. The data has HH and HV polarization and was acquired during dry season in June-July 2007. The SAR data was resampled into 30-m resolution and was registered to LANDSAT ETM image acquired in May 31, 2003. These data were geo-referenced into UTM zone 50 projections and were resampled using nearest neighborhood algorithm, to minimize radiometric changes in the corrected data. The optical data was used as a reference for the AGB and stem volume assessments.

Digital Number (DN) values of each polarization (i.e. HV and HH) were converted into normalized radar backscattering (σ) using following equation:

$$\sigma = 10 \times \left(DN^2 \right) - 83.0 \text{ dB} \quad \text{Eq. 7.2}$$

We used three polarization bands of radar backscattering (σ), namely HH (σ_{HH}), HV (σ_{HV}) and HV/HH ($\sigma_{HV/HH}$) to correlate with the forest attributes. The correlations between the radar backscatter and the forest properties showed non-linear relationship. The above ground biomass was modeled using exponential equation recommended by Luckman, et.al. (1996).

$$\sigma = a \times \left(1 - e^{-bx} \right)^c \quad \text{Eq. 7.3}$$

where x is AGB; a, b and c are free parameters. Non-linear least square method was conducted using Levenberg-Marquardt algorithm to estimate the model which minimizes the error estimate.

Stem volume (V) on the other hand, was modeled using quadratic model,

$$\sigma = a + (b \times V) + (c \times V^2) \quad \text{Eq. 7.4}$$

where a, b, c are equation parameters.

7.3 Results and Discussion

7.3.1 Estimation of AGB and Stem Volume based on field observation data

Multi-linear regression method was used to model stem volume (V) using basal area (BA) as the predictor, and the following equation was generated.

$$V = 0.061 \times BA + 0.494 \quad \text{Eq. 7.5}$$

The utility of basal area to correlate with the stem volume was due to the basal area had strong correlation with stem volume ($r^2 = 0.937$, $df = 13049$, $\alpha = 0.01$, $SEE = 0.28$). We used Eq. 7.5 to estimate stem volume from the sample data collected in 2003.

The AGB estimate was 363.1 ± 143.5 ton/ha (95% of confidence interval) (Tab. 7.1), similar with the finding of Samalca (328 ± 29.7 ton/ha, 95% of confidence interval) who conducted study in a similar forest region for estimating biomass using destructive sampling approach (Samalca, 2007).

Tab. 7.1 Statistics of forest biophysical properties

	DBH (cm)	Basal Area (m ² /ha)	Stem Volume (m ³ /ha)	AGB Estimate (ton/ha)
Mean ($n = 38$)	22.4	31.8	336.8	363.1
Min	14.6	7.4	140.3	65.2
Max	31.8	58.1	545.2	853.0
SD	3.4	10.2	94.5	143.5

We also found that the standard deviation of AGB was relatively high ($\sim 40\%$ of the estimate). This uncertainty is a challenging task for any remote sensing study in predicting the forest biomass, particularly those conducted in tropical forest environments which are characterized by highly complex vegetation structure and uneven forest stand age (Lu, 2006).

For smaller trees ($dbh < 30$ cm), the difference between the AGB and stem volume were considerably high. For larger trees ($dbh > 61$ cm), there was only a slight difference between the stem volume and AGB, as the biomass sharply increased during the regenerating processes. This implies that the functionality of the forests as a carbon sink was highly

disrupted once those forests have been disturbed and the recovery processes to return to the initial state of carbon accumulation, is not an instant process.

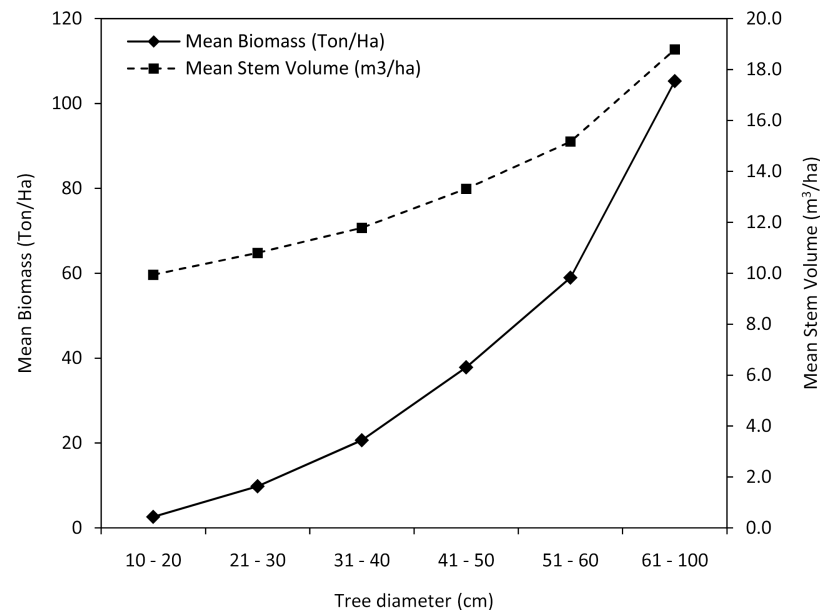


Fig. 7.1 Estimates of above ground biomass and stem volume

7.3.2 Relationship between radar backscatter and forest properties

The HH, HV and HV/HH polarizations of ALOS PALSAR data showed relatively low correlations with the AGB (Fig. 7.2a, Fig. 7.2b, Fig. 7.2c). The HH polarization saturated at -7.5 dB, higher than the HV band that saturated at -15.5 dB. The HV/HH band showed the saturation point at -7.75 dB which was slightly lower than the HH band. Previous study found the HV band has greater dynamic anticipated by microwave interaction models, which make this cross-polarized signal is due predominantly to crown scattering mechanisms and is therefore better coupled to the AGB (Luckman et al., 1996).

Unlike the AGB, the HH band was more sensitive than other bands in explaining the stem volume up to 7.7 dB (Fig. 7.2d). The correlation reversed into a negative direction at higher radar backscatter (Fig. 7.2d). This is unlikely for the radar backscatter, because normally the backscatter values reveal similar correlation characteristics as shown by the correlation with the AGB (Fransson and Israelsson, 1999).

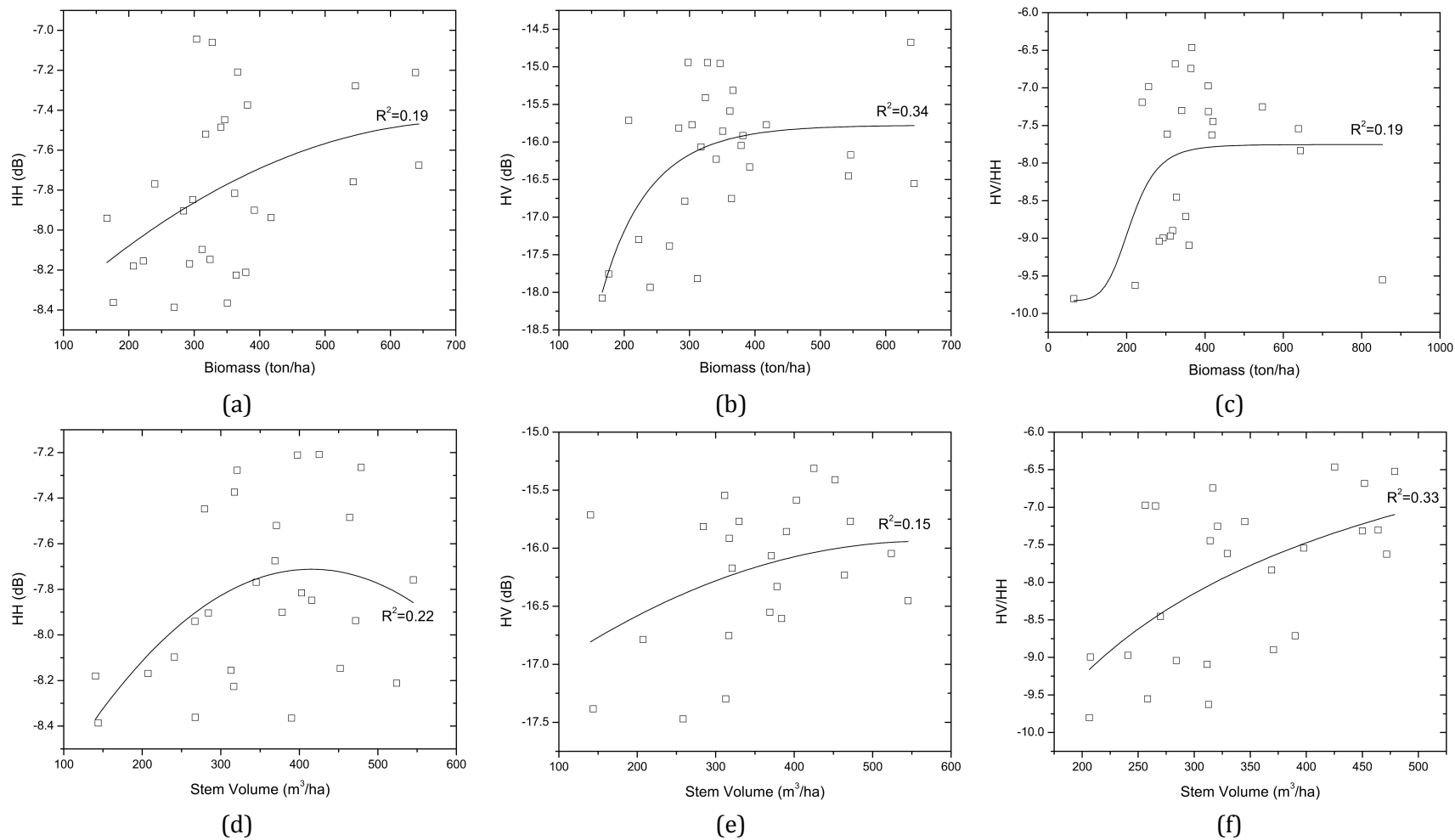


Fig. 7.2 Correlation between (a) HH vs AGB, (b) HV vs AGB, (c) HV/HH vs AGB, (d) HH vs stem volume, HV vs stem volume, and (e) HV/HH vs stem volume

This condition occurred probably due to the sample data being mainly collected on advanced secondary forests which were dominated by complex vegetation and canopy structure. The study area, in fact, is a natural forest dominated with mixed *Dipterocarp* sp., which consists of overstory, canopy layer, understory, and shrub layers. The overstory vegetation may exceed 50 m height, while the understory trees varied from 2 – 5 m. The complexity of the forest structure limited the radar backscatter to partially explain the stem volume variations without the interference from the canopy cover. Another explanation is that the stem volume has lower variations than the AGB that exponentially increased with the tree diameter (Fig. 7.1). Due to this relatively low variation, the radar backscatter was not sensitive to explain higher stem volume, showing negative correlation between both variables.

7.3.3 Radar backscatter modeling

We observed that the HV band could explain the AGB better than other polarization bands ($R^2_{HV}=0.34$ vs $R^2_{HH}=0.19$ and $R^2_{HV/HH}=0.19$) (Fig. 7.2a, Fig. 7.2b, Fig. 7.2c). Thus, the modeling of radar backscatter with the AGB was conducted by fitting the HV band (σ_{HV}). The backscatter – AGB data was fitted and the following model was generated.

$$\sigma_{HV} = -15.776 \times (1 - e^{-(0.0122 \cdot \text{AGB})}) - 0.933 \quad \text{Eq. 7.6}$$

Unlike AGB, the HV/HH band explained the stem volume better than did other polarization bands ($R^2_{HV/HH}=0.33$ vs $R^2_{HH}=0.22$ and $R^2_{HV}=0.15$) (Fig. 7.2a, Fig. 7.2b, Fig. 7.2c). The HV/HH band ($\sigma_{HV/HH}$) – stem volume model is described as follows.

$$\sigma_{HV/HH} = -14.478 + (0.6724 \times V) + (-0.0163 \times V^2) \quad \text{Eq. 7.7}$$

Generally, the modeling of radar backscatter was problematic since the scatter plots of the radar backscatter against the forest properties were greatly disperse (Fig. 7.2). The correlations of these variables, however, were non-linear. Therefore, the predictions of stem volume and AGB from the radar backscatter (σ) were modeled using polynomial approach. The more complex model, e.g. water cloud model, might be used alternatively to better correlate the stem volume with radar backscatter (Fransson and Israelsson, 1999).

7.3.4 Prediction of AGB and Stem Volume using Radar Backscatter

Inversion of the exponential AGB equation (Eq. 7.6) was not successful, as the nature of the modeled variable correlations was changed once the equation was reversed. Instead, a polynomial model was developed to predict the AGB.

$$AGB = -7859 - (1035.258 \times \sigma_{HV}) - (32.635 \times \sigma_{HV}^2) \quad \text{Eq. 7.8}$$

The model might be used to estimate the AGB which was less than 400 ton/ha and the HV backscatter (σ_{HV}) should range from -18.5 dB to -15.75 dB (Fig. 7.3).

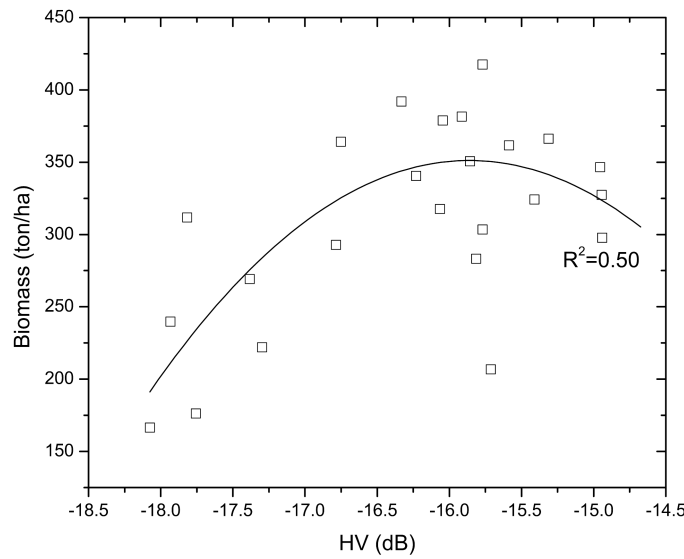


Fig. 7.3 Fitting of AGB model

The AGB estimate was 308.5 ± 52.8 ton/ha ($r=0.35$, $SEE=156.3$) markedly lower than the observed AGB (363.1 ± 143.5 ton/ha). The error estimate was high ($\sim 50\%$ of the estimate) due to high variability of the observed AGB. The estimate, however, might be used as an initial prediction of the AGB density over the forest area, given the mentioned above conditions (i.e. AGB and HV limits) were satisfied.

Stem volume (V) was estimated using the HV/HH backscatter ($\sigma_{HV/HH}$) as predictor resulted in the following polynomial model.

$$V = 256.85 - (65.458 \times \sigma_{HV/HH}) - (6.8 \times \sigma_{HV/HH}^2) \quad \text{Eq. 7.9}$$

Similar with the AGB model, the predictive ability of the stem volume model was rather limited to relatively low values, which were less than 500 m³/ha, and the HV/HH band should range from -10.0 to -6.0 (Fig. 7.4). The estimated stem volume was 340.4 ± 67.1 m³/ha ($r=0.57$, $SEE=67.12$) slightly higher than the observed volume (336.8 ± 94.5 m³/ha).

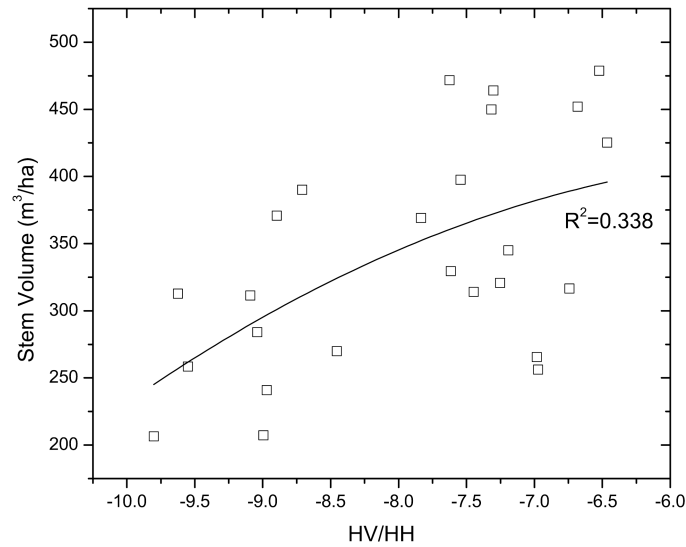


Fig. 7.4 Fitting of stem volume model

7.3.5 Evaluation of AGB Estimate

To facilitate better understanding of the forest properties modeling using mosaic ALOS PALSAR backscatter, the AGB estimate is further discussed in this section.

Forest biomass assessment is important to indicate carbon contents over a forest area, because an accurate mapping of the biomass could explain the amount of absorbed carbon by the forest over certain periods. This study found the estimated AGB density ranged up to 350 ton/ha (Fig. 7.3).

Interpretation of the Landsat ETM image confirmed that lower AGB were found along the rivers and main logging roads, and in the surrounding of settlement areas (see Fig. 7.5). In flat terrains, the AGB estimate resulted in favorable results, but in higher slope areas, like in the southern part of the study site, the estimate was less accurate. This might be probably due to shadow effects from the backside of rugged terrains which returned lower backscatter to the radar sensor (less than -18.5 dB).

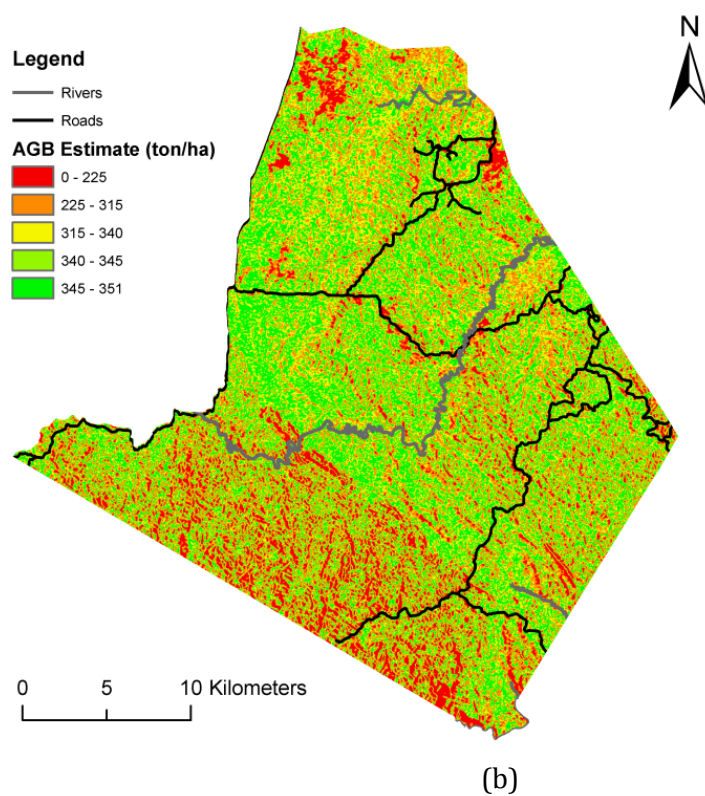
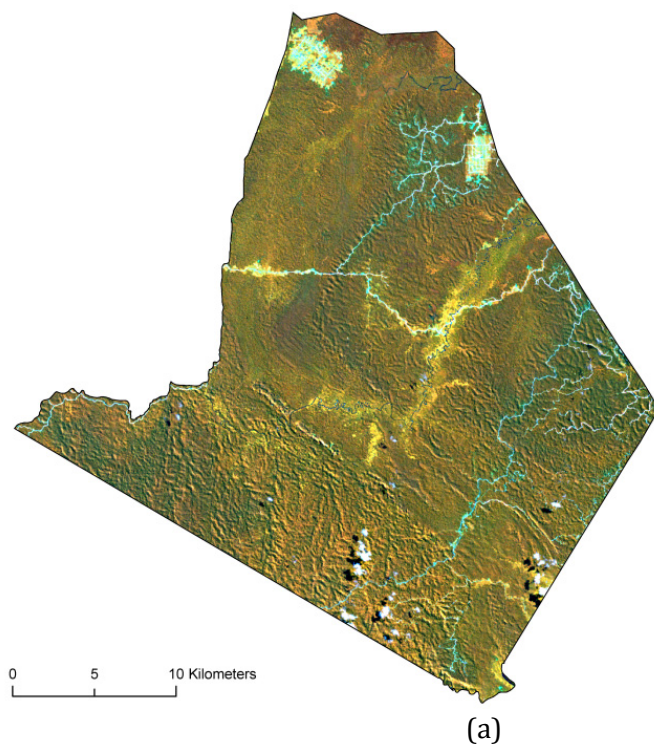


Fig. 7.5 Landsat ETM data with band 4,5,3 in RGB channel (a) and the AGB estimate modeled from HV band (b)

The results might be improved by correcting the topographic effects on the SAR image, but since the image is a mosaic image the topographic corrections could be problematic, especially in finding proper solar azimuth and sun angle values. Although the mosaic SAR data provided by JAXA in K & C Initiatives Project website has already been preprocessed, including for topographic corrections, unfortunately there are still some topographic effects remaining on the image.

The empirical biomass equation modeled from the HV band was sensitive only if the radar backscatter ranged from -18.5 to -15.75; this limitation was rather a problem to accurately map higher AGB values at higher slope regions. Our observation comparing the biomass estimate and the ETM data confirmed that the HV band was more sensitive to assess above ground biomass (AGB) in flat terrains rather than in high slope regions (Fig. 7.5).

The HV band is also more sensitive to estimate the AGB in logged over forests, where even forest stands are found, compared to mature forests or primary forests, which have more variable stand age. The backscattering of mosaic ALOS PALSAR data to characterize volume scattering which usually comes from dense vegetation was partially reduced, leaving only its ability to identify the surface scattering from smooth geometry objects. Therefore, this data could give favorable results to characterize the biomass over regenerating forests, or less complex vegetation landscape, such as shrubs, savanna, and grassland.

The proposed approach may be suitable for regional study estimating forest biomass over a large region where no high estimate accuracy is required. Given the condition that the mosaic ALOS PALSAR data can be downloaded free of charge from Kyoto and Carbon Initiatives Project website (Rosenqvist et al., 2008), this study has demonstrated that the data may be used to estimate the forest attributes from global datasets. JAXA as a provider of the ALOS PALSAR mentioned that the data will always be made available for public and be updated regularly once a year, especially those of tropical regions, such as Indonesia. It means that the data might be alternatively used for regular monitoring of deforestation and forest degradation, as well as for predicting the dynamics of forest biomass over certain period.

7.4 Conclusions

This study found relatively low correlation between radar backscatter and forest properties data, mainly due to high variations of forest properties data under study and the asymptote problem of radar backscatter. The radar backscatter was saturated at higher stem volume and AGB values. Among other bands, the HV polarization band correlated better with the AGB, whereas the stem volume was better explained by the HV/HH band. These estimates,

however, might be used as an initial prediction of the forest parameters, where no high accuracy is needed. The proposed approach has potential for estimating the forest attributes using global datasets. This study has shown that free mosaic data of ALOS PALSAR may be useful for predicting the AGB, and given the availability of multi-temporal datasets the dynamics of biomass density can be monitored regularly.

Chapter 8

Dual-Polarimetry SAR and Optical Remote Sensing Data for Tropical Peatlands Discrimination^{††}

ABSTRACT

This study investigates the potential of dual-polarimetric TerraSAR-X data to characterize different peatlands in tropical swamp forests of Central Indonesia. Discrimination of peatlands is important to assess the carbon contents below the ground. The HH and VV polarizations of TerraSAR-X data are extracted from the Single Look Slant Range Complex (SSC) images. Matrix decomposition is calculated from the covariance matrix resulting in the α -angle, scattering entropy and anisotropy. The images are then converted from slant range into ground range resolution resulting in 6.5 meter of spatial resolution images, and were coregistered with Landsat ETM data. Our results show that scattering entropy/ α -angle plane segmentation of sampled peatlands data were highly overlapped, except for very shallow peat in sparse forest. To estimate the importance of radar backscatter and its polarimetric properties for distinguishing different peatlands, we used a canonical discriminant analysis method. We found that although entropy and anisotropy were strongly correlated, but these variables were more important than HH, VV, and alpha angle for discriminating the peatland classes. Different inputs combining SAR and ETM data were experimented in the maximum likelihood classification, and we found the classification using HH, VV, and entropy of SAR data combined with ETM band 123457 yielded 80% of accuracy. A majority analysis was performed on the classified image, and increased the accuracy up to 87% with Kappa statistics of 0.85.

Keywords: TerraSAR-X, dual polarimetry, peatlands, tropical forests, scattering entropy, α -angle, canonical discriminant analysis, maximum likelihood classification

^{††} This chapter is based on:

Wijaya, A. P.R. Marpu and R. Gloaguen. *Discrimination of Peatlands in Tropical Swamp Forests using Dual-Polarimetric SAR and Landsat ETM Data*. International Journal of Image and Data Fusion, Taylor and Francis Publisher (Accepted for publication).

8.1 Introduction

Peatlands are the most efficient carbon sink sources as compared to all other terrestrial ecosystem (van Beukering et al., 2008). Although peatlands cover only 3% of the Earth land surface, they contain as much carbon as all terrestrial biomass, twice as much as all global forest biomass, and about the same as can be found in the atmosphere. Peatlands store carbon for thousands of years and play critical role in biodiversity conservation and hydrological regulation (Wetlands International, 2007).

There are different types of peatlands, but this study focuses on tropical peatlands, so-called peat swamps. The peat swamps are important ecosystems in Indonesia as they have an abundance of fauna and flora and confirm the uniqueness of the ecosystem and its ecological attributes and values (Rieley and Page, 2005). The very special characteristics of tropical peat swamp arise because it is actually two biological communities – peatland and tropical rain forest – that have evolved together and co-existed for thousands of years.

Indonesia has almost 30 million ha of intact peatlands, which is the largest area in South East Asia and about 7.5% of all peatlands on earth. Kalimantan is one of three regions in Indonesia that has most of the peatland forests (Muhamad and Rieley, 2002). In Kalimantan alone before 1996, there were about 3 million ha of peatlands (Page et al., 2002). Continuous environmental change, such as forest burning and clearing, excessive agricultural lands expansion, and conversion to palm oil plantations has threatened the peat swamp forests stability and makes them prone to forest fires. This was demonstrated during the El Nino event in 1997, where 0.79 Mha of forests in Central Kalimantan had burned (32% of total forests), of which 0.73 Mha of peatlands were burned (92% of total burned forests) (Page et al., 2002). Due to this forest fire event, 0.19 – 0.23 Gt of carbons from the peatlands were released to the atmosphere, and if this number was extrapolated for the entire country, 0.81-2.57 Gt of carbon or equal to 13-40% of mean annual global carbon emission from fossil fuels were liberated in 1997 (Hoekman, 2007).

Below ground carbon contents in the peatlands are estimated from the peat thickness, peat ripening level, and total peatland areas. Usually peat thickness is directly measured on the ground by drilling a peat corer into the peat soil until it reaches the mineral soil layer. In addition the degree of peat ripening and bulk density is also determined directly in the field, and the carbon content in different peat layers and various level of peat ripening are collected from available literature (Wetlands International, 2004). Major problems in classifying peat types and understanding peat-forming processes in the tropics arise from the lack of detailed, standardized and precise site-based information across the equatorial

zone. It is also difficult to accurately determine the boundaries between mineral and peat soils since both support forests, often of similar structure (Rieley and Page, 2005).

Remote sensing is a promising approach for characterizing different types of peatland at large scale, because the data could cover huge areas and eventually it may complement in situ approach which needs a great number of accurate ground truth data. Classification of different peatlands in Kalimantan was successfully carried out using ALOS Palsar data discriminating various types of tropical peat forests (Wielaard and Hoekman, 2009). The potential of Chris Proba data was successfully applied for characterizing different peatlands in similar forest landscape (Liesenbergh et al., 2009). The present study, however, proposes the utility of TerraSAR X image to characterize different peat thickness associated with the peatland types above the ground. The peat thickness in Kalimantan varies from very shallow (<50 cm) to very deep layers (up to 20 m) (Rieley and Page, 2005). The identification of peat thickness using satellite imagery is rather impossible due to the complexity of peat soils composition and the vegetation structure on the top of soils surface, which could have similar vegetation structure for different peat types and thickness. Therefore, relative association between certain peatlands and peat thickness should be considered.

This study evaluates the potential of dual polarimetric high resolution SAR data for mapping different peatland forests. Variations of HH and VV polarization and α -angle (alpha), scattering entropy (H) and anisotropy (A) layers generated from the decomposition of covariance matrix were used. We generated the scattering entropy/ α -angle plane segmentation and discussed the results. Additionally, canonical discriminant analysis was used to determine the importance of predefined variables of the SAR data for distinguishing the peatlands. Image classification was carried out using variations of the SAR data afterwards. Combination with the optical Landsat ETM data was also experimented to improve the classification accuracy. Contributions of the dual-polarimetric SAR data for mapping the peatland forests were then discussed.

8.2 Data and Methods

8.2.1 Preparation of satellite data and ground truth data collection

Dual-polarimetric TerraSAR X data used for this study was acquired on March 13, 2008. The HH and VV polarization bands were extracted from the Single Look Slant Range Complex (SSC) images and converted into 54-looks complex images to enhance the radiometric resolution. The enhanced Lee filter was applied to reduce speckle noise and to improve the visual characteristics of the SAR data. Matrix decomposition is generated from the covariance matrix calculating the α -angle (alpha), scattering entropy (H) and anisotropy (A).

The SAR images were converted from slant range into ground range resolution resulting in 6.5 meter of spatial resolution images, and georeferenced with the Landsat ETM data acquired on August 19, 2004, and resampled into 30 meter resolution using nearest neighborhood algorithm. The images were projected into WGS 1984 coordinate system and UTM Zone 49 datum (Fig. 8.1).

Peatlands distribution map and below ground carbon contents were acquired from Wetlands International – Indonesian Program, which conducted study for mapping and estimating the carbon content in Indonesian peatland forests from 2000 – 2002 (Wetlands International, 2007). Land cover map over the study area was acquired from National Coordination Agency for Surveys and Mapping (Bakosurtanal). This map was basically generated using a 1:50,000 topographic map and visual interpretation of the Landsat data series from 2001 – 2003. These maps were overlaid on the top of the orthorectified SAR data and used as a basis for collecting the ground truth data. Polygons of sample data identified four peatland classes, i.e. shallow peat in secondary forest (PSS), very shallow peat in sparse forest (PVSp), deep peat in primary forest (PDP), shallow peat in primary forest (PSP). We also collected sample data from mangrove forest (MF) areas as this forest shares similar characteristic as the peatlands, but contains no peat soils, and is also included as wetlands (Tab. 8.1). Additionally, for maximum likelihood classification combining SAR and Landsat ETM data, samples of cultivated lands (CL), palm oil plantation (PO), non-vegetated areas (NV) and water body (Wt) were also collected.

8.3 Methods

8.3.1 Matrix decomposition of SAR data

The H/A/Alpha polarimetric decomposition is based on an eigenvector decomposition of the (2×2) complex covariance [C2] matrix,

$$\langle [C2] \rangle = [V] \cdot [\Lambda] \cdot [V]^{-1} \quad \text{Eq. 8.1}$$

where [Λ] and [V] represent (2×2) real eigenvalue and special unitary eigenvector matrices, respectively.

The (2×2) complex covariance matrix [C2] is a hermitian and semi-definite positive matrix, and its eigenvectors are orthogonal and the respective eigenvalues are real positive. The eigenvector decomposition of certain distributed target covariance matrix is considered as a simple statistical model consisting in the expansion of the (2×2) complex covariance matrix into a weighted sum of two covariance matrices (European Space Agency (ESA), 2007),

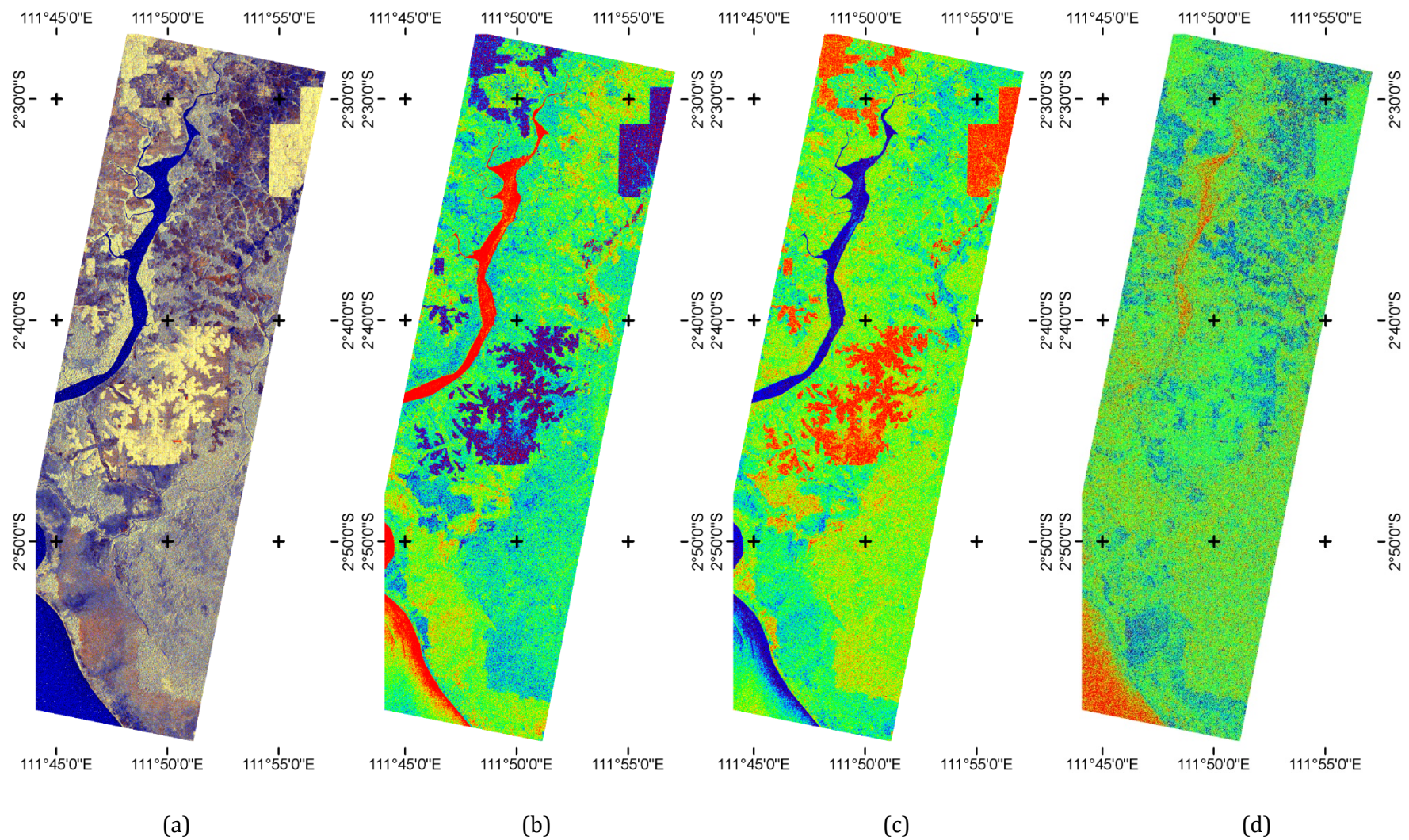


Fig. 8.1 Dual polarimetry SAR data: (a) RGB combination of HH, VV, HH-VV bands, (b) scattering entropy, (c) anisotropy, and (d) alpha angle

Tab. 8.1 Characteristics and carbon contents below the ground of different peatlands found in the study area

Peat classes	Peat types/ Proportions (%)	Peat thickness	Peat carbon stocks (ton/ha)	Land cover type
Mangrove forest (MF)	-	-	-	Mangrove forest
Deep peat in primary swamp forest (PDP)	Hermists/fibrists (H3a)/(60/40)	2 – 4 m (deep)	1158 ton/ha	Primary forest
Shallow peat in primary swamp forest (PSP)	Hermists/fibrists/ mineral (H1b) / (50/30/20)	0.5 – 1 m (shallow)	232 ton/ha	Primary forest
Very shallow peat in sparse forest (PVSp)	Hermists-mineral (H1i) / (20/80)	<0.5 m (very shallow)	41.5 ton/ha	Sparse forest
Shallow peat in secondary swamp forest (PSS)	Hermists/fibrists/ mineral (H1b) / (50/30/20)	0.5 – 1 m (shallow)	232 ton/ha	Secondary forest

Source: (Wahyunto et al., 2004)

$$\langle [C2] \rangle = \sum_{i=1}^2 \lambda_i v_i v_i^{T*} = \lambda_1 [C2]_1 + \lambda_2 [C2]_2 \quad \text{Eq. 8.2}$$

Pseudo-probabilities of the (2×2) complex covariance matrix expansion element are defined from the set of sorted eigenvalues.

$$p_i = \frac{\lambda_i}{\sum_{j=1}^2 \lambda_j} = \frac{\lambda_i}{span}, \text{ with } p_1 \geq p_2, \quad \text{Eq. 8.3}$$

where p_i is the pseudo-probability of the eigenvalue λ_i , representing the relative importance of this eigenvalue with respect to the total scattered backscatter power, and by convention $p_1 \geq p_2$. The scattering entropy (H) is calculated using the following formula,

$$H = -p_1 \log_2(p_1) - p_2 \log_2(p_2) \quad \text{Eq. 8.4}$$

The anisotropy (A), which may also be used to characterize the scattering mechanism of peatlands was calculated as follows.

$$A = \frac{p_1 - p_2}{p_1 + p_2} \quad \text{Eq. 8.5}$$

The distribution of the scattering mechanism probabilities can be fully described by the entropy (H) and anisotropy (A). The Entropy indicates degree of statistical disorder of the scattering phenomenon, whereas the anisotropy is defined as the relative importance of the secondary scattering mechanisms. For high entropy values, i.e. more than 0.7, the anisotropy is used to fully characterize the set of scattering probabilities.

Additionally, mean alpha angle (α -angle) is also generated from

$$\alpha = \sum_{i=1}^2 p_i \alpha_i \quad \text{Eq. 8.6}$$

The alpha angle may be interpreted as follows: $\alpha \sim 0$ means that the scattering mechanism corresponds to the single bounce scattering, mainly from a rough surface, $\alpha \sim 45^\circ$, means that the volume scattering dominates the scattering mechanism, and if $\alpha \sim 90^\circ$ the scattering mechanism is mainly due to double-bounce scattering.

8.3.2 Image classification

The HH, VV, entropy, anisotropy and alpha angle were used to model the peatlands data under study (Tab. 8.1). Canonical discriminant analysis is conducted to maximize the

difference between values of different peatland classes, and to determine the most weighted variable in discriminating the peatland class. Correlation between the predictors was also estimated following Pearson correlation procedure. The classification using discriminant analysis (Liesenberg et al., 2009) was then carried out and the classification accuracy was assessed from cross-validation dataset using confusion matrices. We also experimented with more robust classification technique based on Bayesian approach, maximum likelihood technique, combining different variations of input data (i.e. HH, VV, entropy, anisotropy, and alpha). Also, integration with the optical Landsat ETM data was experimented and similarly the classification accuracy was assessed using confusion matrices. The observation on visual characteristics of the classified image was evaluated and the results were discussed.

8.4 Results and discussion

8.4.1 SAR Backscatter Responses

Radar backscatter responses from the HH and VV bands were relatively high for mangrove forests (MF) (> -2.0 dB) and low for very shallow peat in sparse forests (PVSp) class (< -3 dB) (see Fig. 8.2). This is definitely due to different dominant scattering mechanism between both classes. More homogenous vegetation structure and top canopy height were found in mangrove forests than in the sparse forest.

There was a slight difference between the mean HH and VV values of the PDP, PSP and PSS classes. This was because the high overlap between the radar backscatter of the primary and secondary forests, therefore the utility of the HH and VV bands solely for the discrimination of peat types under these landscape might be problematic. In fact, the correlation between the HH and VV was relatively strong ($r = 0.722$, p -level 0.01, two-tailed test) as compared to the correlations between these bands with the entropy, anisotropy and alpha angle ($r < 0.30$).

For each peatland class, anisotropy showed almost exact mirrored values to that of the scattering entropy (Fig. 8.2). The anisotropy values increased with the peat soil thickness and vice versa for the entropy. Pearson coefficient correlation confirmed that both variables were highly correlated ($r = -0.991$, p -level 0.01, two-tailed test). The alpha angle (α -angle), however, revealed similar trends as the radar backscatter in overall, and it might still be useful to improve the discrimination of the peatland types.

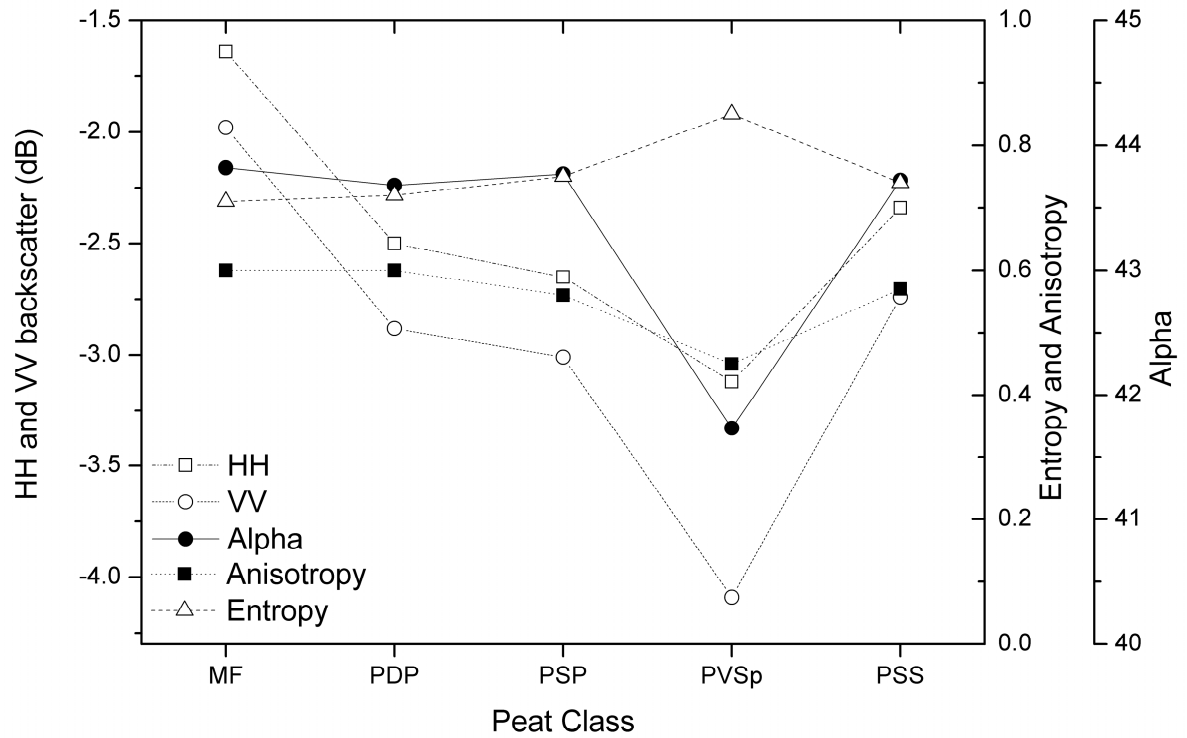


Fig. 8.2 Radar backscatter, entropy, anisotropy, and alpha angle responses to different peatland classes

The α -angle/scattering entropy plane segmentation is plotted and we found that PSS, PDP and PSP were highly overlapped (Fig. 8.3). These classes, including the PVSp class, show dominantly volume scattering mechanism (segmentation plane area 4), where the alpha values are moderate and the entropy of over 0.5. The PVSp class was partly overlapped with other peatland classes, due to the occurrence of surface scattering mechanism in this particular class, as an effect of more even forest canopy and less complex vegetation structure.

Dual polarimetry TerraSAR X data is rather restricted for characterizing the thickness of peat soils, because of the limitation of x-band in penetrating through canopies over primary and secondary forests, i.e. the thickness of shallow and deep peat soils under these forest canopies could not be easily discriminated. However, in sparse forest or less vegetated areas, we can expect that this SAR data is useful for the peatland classification.

8.4.2 Canonical Discriminant Analysis

To provide better understanding on the importance of radar backscatter and matrix decomposition features for the classification of peatlands, the canonical discriminant analysis was applied. The F-test indicates that the homogeneity assumption of the covariance matrices has not been met (approx. F-test value = 703.58, p -level < 0.0001).

However, this test is strongly influenced by data non-normality and may not be accurate. This is a general case, especially for remote sensing study which is dealing with high dimensional data.

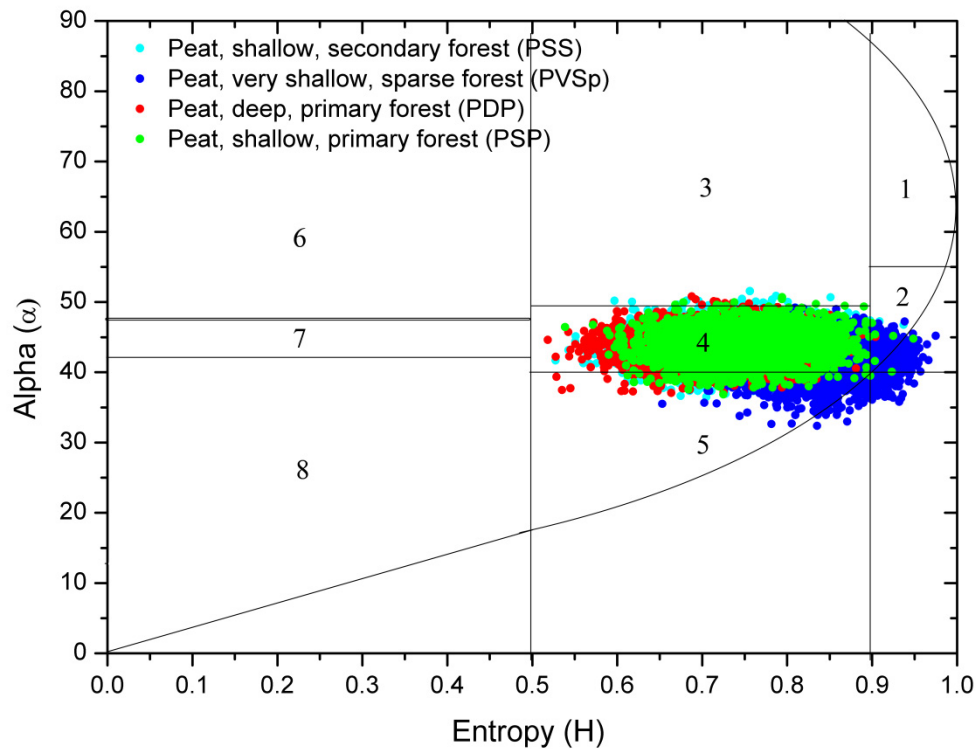


Fig. 8.3 Sample data plotted on the alpha angle / scattering entropy segmentation plane. In general, regions 1, 3 and 6 identify regions are dominated by multiple scattering, regions 2, 4 and 7 are dominated by volume scattering, and regions 5 and 8 characterize surface scattering mechanism

Tab. 8.2 Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	p-value.
Alpha	.869	821.628	4	21752	.000
Anisotropy	.535	4731.239	4	21752	.000
Entropy	.570	4109.281	4	21752	.000
HH band	.709	2235.333	4	21752	.000
VV band	.607	3523.479	4	21752	.000

The discriminant analysis was conducted to assess whether the HH, VV, entropy, anisotropy and alpha angle could correctly distinguish the peatlands under study. The F-test for equality of group means confirmed that each predictor was individually significant to differentiate the peatlands (p -value < 0.0001) (Tab. 8.2).

Four empirical functions were generated from the canonical discriminant analysis and used for maximizing the differences of the peatland classes. Wilk's lambda tests (not shown here) was significant for each model combinations (p -value < 0.0001), except if the model 4 was solely tested (p -value > 0.05). The absolute correlation between each variable and any discriminant function revealed that anisotropy and entropy were strongly correlated with the function 1 (r_{anis} -0.559, r_{ent} = 0.67) and 3 (r_{anis} 0.77, r_{ent} = -0.716), whereas the HH (r = -0.795) and VV (r = -0.684) bands had strong correlations with the model 2. The alpha angle, on the other hand, was highly correlated with the empirical model 4 (r = -0.875). The standardized canonical discriminant function coefficients suggested that the entropy and anisotropy contributed most to distinguishing the peatlands. However, both variables were also highly correlated, so that the joint use of both variables might be redundant and might reduce the predictive ability of any classification methods, especially if the classification algorithm works based on some statistical assumptions.

The classification results show that the models had favorable classification accuracy merely for the PVSp (78.3%) and MF (57.6%) classes (Tab. 8.3). Due to complex vegetation structure, the accuracy of the PSS, PDP and PSP classes were considerably low (< 50%). The highly overlapped radar backscatter values of these classes might explain low classification accuracy results (Fig. 8.4). In overall the accuracy of discriminant analysis classification was 46.3%, which was unfavorable. Therefore, we used Maximum likelihood technique, which is well-known for its robustness to classify mixed land cover classes (Mather, 2004).

Tab. 8.3 Classification results and accuracy assessment based on cross validation ^(a,b)

	Peatland Class	Predicted Group Membership					Total
		PSS	PVSp	PDP	PSP	MF	
Count	PSS	1324	333	1210	1738	1252	5857
	PVSp	14	3735	36	985	0	4770
	PDP	537	99	1313	1007	844	3800
	PSP	734	312	935	1914	344	4239
	MF	383	42	539	348	1779	3091
%	PSS	22.6	5.7	20.7	29.7	21.4	100.0
	PVSp	.3	78.3	.8	20.6	.0	100.0
	PDP	14.1	2.6	34.6	26.5	22.2	100.0
	PSP	17.3	7.4	22.1	45.2	8.1	100.0
	MF	12.4	1.4	17.4	11.3	57.6	100.0

^aCross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

^b46.3% of cross-validated grouped cases correctly classified.

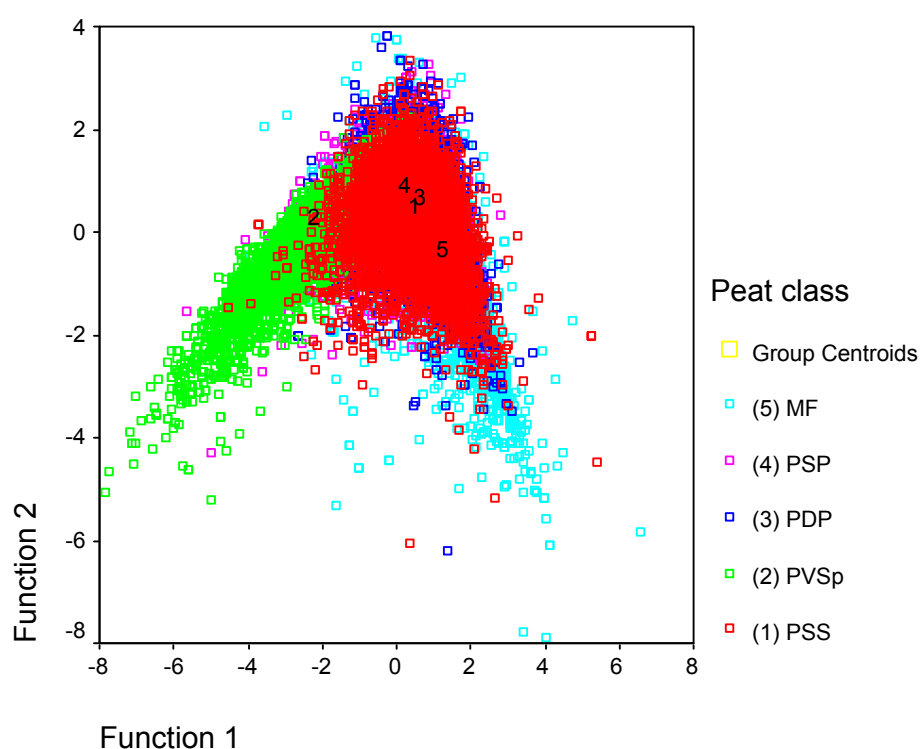


Fig. 8.4 Canonical discriminant function plot. The function 1 and 2 refer to the functions generated using HH, VV, entropy, anisotropy, and alpha as predictors, employing the coefficient of discriminant analysis

8.4.3 Classification of Radar Backscatter and Optical Data

Given the discriminant analysis results, we experimented with supervised maximum likelihood classification combining HH and VV bands, and SAR polarimetric properties, i.e. alpha angle, entropy and anisotropy for peatlands classification. The individual use of radar data found that the combination of HH, VV, alpha and entropy yielded higher accuracy (overall accuracy 49%, Kappa 0.43) than other SAR data combinations (Fig. 8.5). The confusion matrices show that the dual-polarimetric SAR data was suffered for discriminating the peatlands in primary and secondary forest. This is because the polarimetric signatures of volume scattering in X-bands is somehow limited and could only partially explain the scattering phenomenon over the forest areas. Moreover, the maximum likelihood algorithm has similar problem as the discriminant analysis method to accurately classify the PSS, PDP and PSP classes using only radar backscatter inputs.

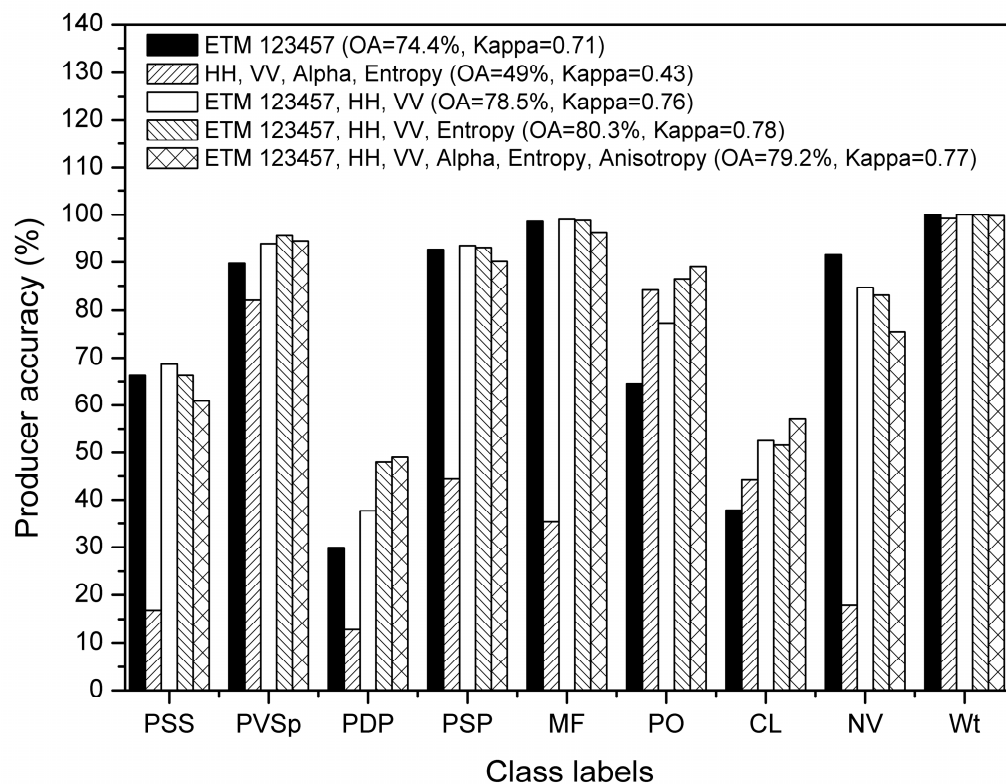


Fig. 8.5 Comparison of classification accuracy using confusion matrices for each class label, namely shallow peat – secondary forest (PSS), very shallow peat – sparse forest (PVSp), deep peat – primary forest (PDP), shallow peat – primary forest (PSP), palm oil plantation (PO), cultivated lands (CL), non-vegetated lands (NV) and water body (Wt)

Combining the SAR backscatter with optical ETM data has greatly increased the accuracy up to 80% using the HH, VV, entropy and ETM band 123457 in the classification (Fig. 8.5). An attempt to combine entropy and anisotropy of SAR data with ETM bands slightly reduced

the classification accuracy, because of high correlation of both SAR features as observed from the discriminant analysis results. Fusion of both sensors for the classification has overcome the limitation of individual use of SAR data for discriminating highly overlapped training signatures on PSS, PDP, and PSP classes. However, a large difference in spatial resolution (SAR data has originally 6.5 m and Landsat ETM is 30 m of spatial resolution) and different properties acquired by both sensors caused the classification arrived at mixing results in some extent. Radar backscatter captured geometric properties of objects on the ground, whereas the ETM surface reflectance measured the energy radiated back from the ground objects.

Fig. 8.6 showed that due to higher spatial resolution and the advantage of different scattering mechanism characteristics from smooth and rough objects on the ground (e.g. non-vegetated and dense vegetation), the SAR data classification produced relatively more accurate delineation than the ETM data in less complex and sparse vegetation, but overlooked the peatlands classification on primary and secondary forests (Fig. 8.6b). The ETM data, on the other hand, yields more accurate results than the SAR data for characterizing the mangrove forest and in discriminating the peatlands on primary and secondary forests (Fig. 8.6a). The fusion of the SAR backscatter and ETM image (Fig. 8.6c and Fig. 8.6d) could resolve this problem, although the classifications of PSS and PDP classes were still problematic (Tab. 8.4). To improve the classification result, we filtered the classified image using a majority analysis employing 3×3 moving window to minimize minority spurious pixels within certain class labels. This process significantly increases the classification accuracy (overall accuracy 87%, kappa 0.85) and the visual appearance of the classified image as well (see Fig. 8.6d and Tab. 8.4).

Better classification results could be achieved employing longer wavelength of SAR image (e.g. ALOS Palsar with L-band) which could penetrate forest canopy, or by applying better filtering processes prior to the classification. The SAR data has more advantage than the ETM data due to its cloud penetration capability, and especially for the study in tropical humid regions it may have great benefits given the data availability.

8.5 Conclusion

Contribution of dual-polarimetric SAR data to characterize tropical peatlands was investigated in this study. The HH, VV, entropy, anisotropy and mean alpha angle might be useful for discriminating the peat types under some limitation. Discriminant analysis method revealed that anisotropy and entropy were more important for peatlands classification than other SAR features (i.e. HH, VV and alpha angle). However, this method

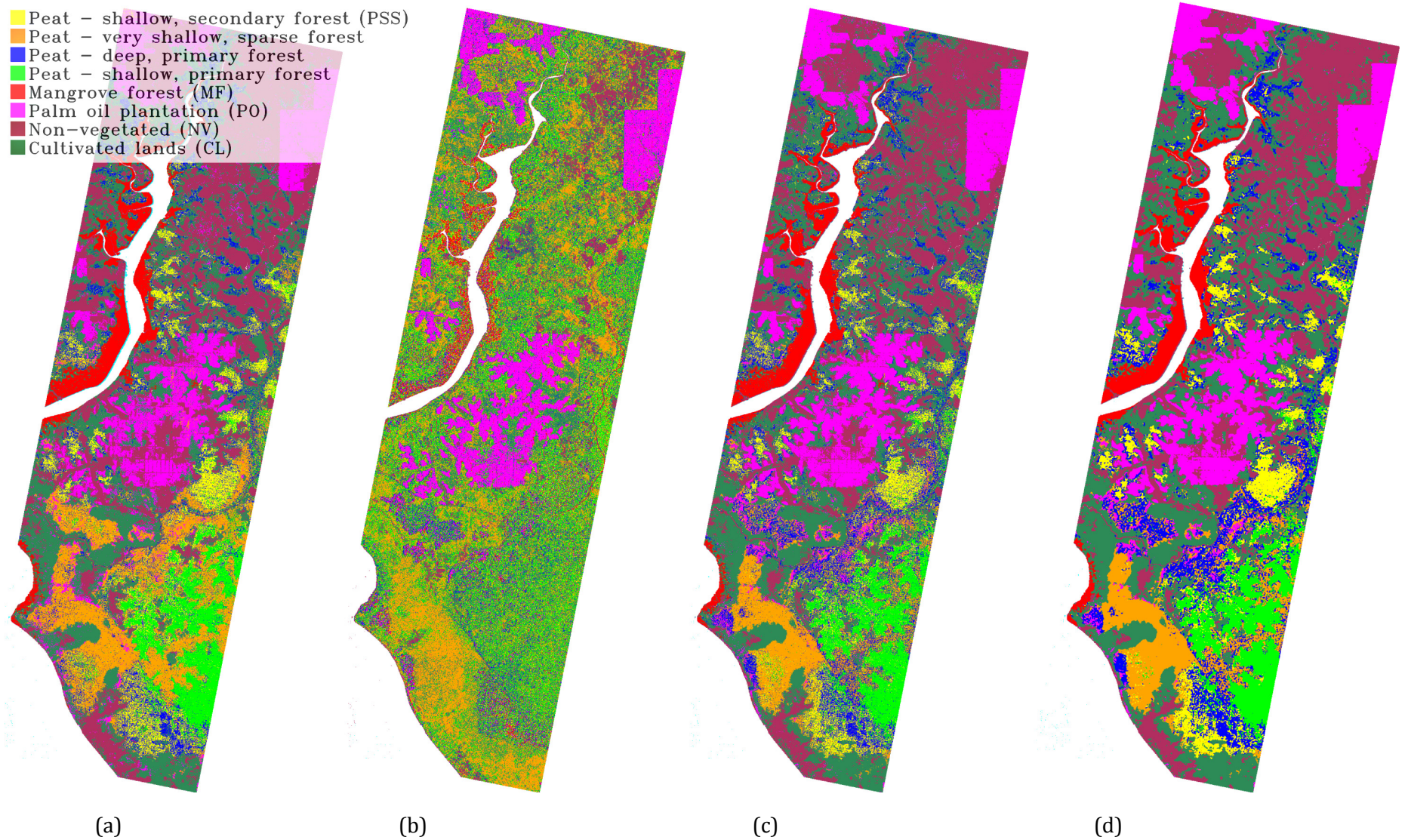


Fig. 8.6 Maximum likelihood classification: (a) ETM Band 123457, (b) HH, VV, Alpha, Entropy, (c) ETM Band 123457, HH, VV, Entropy and (d) ETM Band 123457, HH, VV, Entropy-filtered with majority analysis. The water body (Wt) class label is masked out for better visual representation

Tab. 8.4 Accuracy assessment of the classification using ETM band 123457, HH, VV and entropy, and filtered with majority analysis using 3×3 moving window*

Class labels	PSS	PVSp	PDP	PSP	MF	Wt	PO	NV	CL	User Accuracy (%)
PSS	439	0	92	0	0	0	0	0	52	75.3
PVSp	0	504	0	2	0	0	0	0	0	99.6
PDP	63	0	301	2	0	0	0	0	53	71.8
PSP	1	0	98	473	0	0	0	0	0	82.7
MF	0	0	0	0	500	0	0	0	0	100.0
Wt	0	0	0	0	0	507	0	0	0	100.0
PO	0	0	0	0	0	0	499	8	0	98.4
NV	0	0	0	0	3	0	48	466	141	70.8
CL	1	2	0	0	0	0	0	34	272	75.3
Producer Accuracy (%)	87.1	99.6	61.3	99.2	99.4	100.0	91.2	91.7	52.5	

* Overall accuracy = 87%, Kappa statistics = 0.85

found unfavorable results and yielded 46.3% of accuracy for discriminating the peatland classes.

Using a supervised maximum likelihood classification, we found the classification of HH, VV, and entropy of SAR data combined with ETM band 123457 yielded 80% of accuracy. A majority analysis was performed on the classified image, and the classification accuracy increased up to 87% with Kappa statistics of 0.85.

The fusion of the SAR backscatter and ETM image has improved the classification result, but the classification of shallow and deep peat soils under relatively closed canopy (i.e. primary and secondary forests) is still problematic.

8.6 Acknowledgement

The authors would like to acknowledge the Infoterra GmbH for providing the dual-polarimetric TerraSAR X data for the purpose of this study. The Landsat ETM data was downloaded from the USGS website, and the image is a product from the global land survey (GLS) project.

Chapter 9

Summary

9.1 Thesis Contributions

This thesis mainly focuses on two subjects related to the application of remote sensing data for land cover classification using non-parametric methods (i.e. neural networks and support vector machine methods) and development of a non-destructive approach for above ground biomass (AGB) and forest attributes estimation employing multi-source remote sensing data. Information provided by reliable land cover map is useful for management of forest resources and to support sustainable forest management. Generation of a non-destructive procedure to model forest biophysical properties (e.g. biomass and stem volume), on the other hand, is required to assess the forest resources more efficiently, and coupled with remote sensing data the model can be applied over large areas. Contributions of the present work include:

1. Improved strategy for land cover mapping and above ground biomass estimation using multi-source, multi-resolution remote sensing data.
2. Development of a non-destructive approach for estimating forest biophysical properties applying non-linear analysis based on neural network algorithm.
3. Generation and inclusion of Geostatistics texture features in spectral data classification to provide reliable information provided by a land cover map.
4. Analysis on different features of optical and SAR data for tropical peatlands classification, also fusion of both sensors to improve the discrimination of the peatlands.

9.2 Research Questions Answered

According to the results obtained from this study the research questions formulated in chapter 1 are answered as follows.

1. Which texture features are useful to improve the classification accuracy over tropical forest landscape?

Chapter 3 of the thesis presents the classification of complex forest landscape using Geostatistics texture features. The results concluded that the texture features generated using madogram, rodogram and fractal dimension provide useful information which can

increase the accuracy of spectral data classification. The accuracy increases up to 9% combining the spectral data with texture features in the classification (i.e. overall accuracy of maximum likelihood classification using spectral data only is 79.3%, while combination with texture features inflates the accuracy up to 85.4%). Additionally, maximum likelihood method outperforms support vector machine (SVM) and MLP neural network approaches resulting in higher classification accuracy.

2. How multi-source remote sensing data coupled with in situ data could improve the assessment of above ground biomass (AGB)?

This study proposes a non-destructive approach for above ground biomass (AGB) combining remote sensing and number of stems. Based on the results of chapter 4 higher correlation between estimated and observed AGB is obtained if the reflectance values, vegetation indices, DEM and texture features are included in the AGB modeling. Additionally, application of non-parametric regression method based on Levenberg Marquardt neural network algorithm yields lower error estimates than does of statistical multi regression method.

3. Which types of remotely sensed data and vegetation indices are important for the modeling of AGB?

In chapter 5, spectral data of Landsat ETM, vegetation indices, image transform layers, simple ratios, PCA, tasseled caps bands, GLCM texture features and DEM were generated and combined to model AGB and stem volume. The GLCM mean texture has higher correlation coefficient (r) than other remote sensing (RS) data. Chapter 6 of the thesis explores the capability of bi-temporal Landsat data for estimating basal area, stem volume and AGB. Combining remote sensing and in-situ data, the estimated model can be applied to predict the biophysical properties of different stages of successional forests resulting in favorable error estimates.

4. How land cover changes could affect the dynamics of forest biomass?

This study found that estimated above ground biomass increased with respect to forest successional stage and/or the complexity of vegetation structure. We defined mature forests or advanced secondary forests as the most complex vegetation structure and shrubs as the least complex structure. Stem volume and biomass are markedly increased during the regeneration stages and reached relatively stable condition in advanced/mature forest stage. This is confirmed in chapter 5 which found similar biomass density over mature forest-hilly and very dense forest-hilly landscape (Fig. 5.4).

5. How to estimate the AGB changes using bi-temporal satellite data? What is the predictive ability of linear regression method as compared to a non-parametric neural network model?

Upon radiometric calibration of multi-dates satellite images, the biomass model developed in one image scene can be applied for other image scenes. Unfortunately, problems due to atmospheric variations and attenuations may still persist even if the images are atmospherically corrected and radiometrically calibrated using any standard methods. In chapter 6, multivariate alteration detection technique was applied for radiometric calibration of multi-date Landsat ETM images. This method works based on statistical measures and parameters to calibrate histogram of the target image to the values of reference image. The calibrated target image and the reference image were statistically compared, and we found histograms of both images are not significantly different. Thus, the biomass model generated from the reference image is applied to the normalized image scene. We also tested the ability of multiple linear regression method and neural network model to generate basal area, stem volume and biomass models. The results show that the neural network model has lower error estimates in predicting the forest parameters as compared to the statistical linear regression approach.

6. Which features generated from Synthetic Aperture Radar (SAR) data is more important to assess the biomass and stem volume? How accurate the estimated biomass model?

The modeling of biomass and stem volume using normalized radar backscatter generated from mosaic ALOS Palsar data concluded that HV band is suitable for predicting the biomass while HV/HH band is useful for stem volume estimation. The results, however, suggested that the models are favorable for initial assessment of biomass and stem volume over forested landscape, due to high error estimate and low correlation between SAR data and AGB and stem volume mainly affected by saturation of the SAR data.

7. How the combination of Landsat ETM sensor and SAR data improves the classification of peatland forest?

Chapter 8 investigates the possibility of combining Landsat ETM data (optical sensor) and TerraSAR X data (Microwave sensor) for discriminating peatlands over tropical forest environment. The optical data measures mainly physical properties of ground objects, whereas the microwave sensor captures geometric properties of the ground objects. Combination of both sensors can be an advantage to achieve better classification

results. Using a supervised maximum likelihood classification, we found the classification of HH, VV, and entropy of SAR data combined with ETM band 123457 yielded 80% of accuracy. A majority analysis was performed on the classified image, and the classification accuracy increased up to 87% with Kappa statistics of 0.85. Although the fusion of SAR backscatter and ETM image has improved the classification results, but the classification of shallow and deep peat soils under relatively closed canopy (i.e. primary and secondary forests) is still problematic.

9.3 Conclusion and Recommendations

The final conclusion and recommendations of this study are briefly described as follows.

1. The inclusion of spatial information based on Geostatistics texture feature improves the classification of ETM data over complex forested landscape.
2. Modeling of AGB and stem volume provides more accurate results when remote sensing and in-situ measurement data (based on non-destructive sampling) are combined.
3. Combining multi-source remote sensing data, neural network method has higher accuracy than statistical multi-linear regression technique in modeling the forest parameters.
4. Biomass model estimated from a satellite image can be applied to another image scene upon proper radiometric calibration.
5. Application of Mosaic ALOS Palsar data to predict biomass and stem volume yields favorable estimates which are sufficient for initial assessment of both forest properties.
6. Fusion of optical sensor and SAR data has reduced penetration ability problem of each sensor and improved the classification of tropical peatlands.
7. For future study we recommend that:
 - a. better image classification results may be achieved by combining texture features of SAR data with spectral information from optical data.
 - b. more accurate biomass estimate and forest parameters modeling can be obtained using laser data (LIDAR) or Polarimetric radar interferometry (PolinSAR) data.
 - c. the size of sampling plots on the ground should represent variations of forest stand parameters under study taking into account spatial resolution of satellite data being used.

References

- Anaya, J.A., Chuvieco, E. and Palacios-Orueta, A., 2009. Aboveground biomass assessment in Colombia: A remote sensing approach. *Forest Ecology and Management*, 257(4): 1237-1246.
- Ardo, J., 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, 13(9): 1779-1786.
- Ardo, J., Pilesjo, P. and Skidmore, A., 1997. Neural networks, multitemporal landsat thematic mapper data and topographic data to classify forest damages in the Czech republic. *Canadian Journal of Remote Sensing*, 23(3): 217-229.
- Ashton, H., 2006. *McGraw-Hill Concise Encyclopedia of Bioscience*, 20(1). McGraw-Hill Education, 972 pp.
- Asner, G.P., Keller, M., Pereira, R. and Zweede, J.C., 2002. Remote sensing of selective logging in Amazonia: Assessing limitations based on detailed field observations, Landsat ETM+, and textural analysis. *Remote Sensing of Environment*, 80(3): 483-496.
- Atkinson, P.M. and Tatnall, A.R.L., 1997. Introduction Neural networks in remote sensing. *International Journal of Remote Sensing*, 18(4): 699-709.
- Austin, J., Mackey, B. and Van Niel, K., 2003. Estimating forest biomass using satellite radar: an exploratory study in a temperate Australian Eucalyptus forest. *Forest Ecology & Management*, 176(1-3): 575 - 583.
- Barbosa, P.M., Stroppiana, D., Gregoire, J.M. and Pereira, J.M.C., 1999. An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions. *Global Biogeochemical Cycles*, 13(4): 933-950.
- Bazi, Y. and Melgani, F., 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(11): 3374-3385.
- Berry, W.D. and Feldman, S., 1985. *Multiple Regression in Practice*. Sage University Paper Series on Quantitative Applications in the Social Sciences (series no. 07-050). Sage, Newbury Park, CA.
- BFMP, 1997. The Overall Workplan of Berau Forest Management Project (BFMP), Contract No.: B7 – 5041 / I / 9421, Berau Project Management Unit (BMU), Jakarta.
- Bischoff, W. et al., 2005. Secondary succession and dipterocarp recruitment in Bornean rain forest after logging. *Forest Ecology and Management*, 218(1-3): 174-192.

- Biswasa, M.K., Ghoseb, T., Guhab, S. and Biswasa, P.K., 1998. Fractal dimension estimation for texture images: A parallel approach. *Pattern Recognition Letters*, 19(3-4): 309-313.
- Boyd, D.S., Foody, G.M. and Ripple, W.J., 2002. Evaluation of approaches for forest cover estimation in the Pacific Northwest, USA, using remote sensing. *Applied Geography*, 22(4): 375-392.
- Brown, I.F. et al., 1995. Uncertainty in the biomass of Amazonian forests: An example from Rondônia, Brazil *Forest Ecology & Management*, 75(1-3): 175-189.
- Brown, S., 1997. Estimating Biomass and Biomass Change of Tropical Forests: a Primer. FAO Forestry Paper - 134. FAO - Food and Agriculture Organization of the United Nations, Rome.
- Brown, S. and Gaston, G., 1995. Use of forest inventories and geographic information systems to estimate biomass density of tropical forests: Application to tropical Africa. *Environmental Monitoring and Assessment*, 38(2-3): 157-168.
- Brown, S., Gillespie, A.J.R. and Lugo, A.E., 1989. Biomass estimation methods for tropical forests with applications to forest inventory data. *Forest Science*, 35(4): 881-902.
- Brown, S., Iverson, L.R. and Lugo, A.E., 1994. Land-use and biomass changes of forests in Peninsular Malaysia from 1972 to 1982: GIS approach. *Effects of Land-use Change on Atmospheric CO₂ Concentrations: South and Southeast Asia as a Case Study*. Springer, New York, pp. 117-143 pp.
- Brown, S. and Lugo, A.E., 1984. Biomass of tropical forests - a new estimate based on forest volumes. *Science*, 223(4642): 1290-1293.
- Brown, S. and Lugo, A.E., 1992. Aboveground biomass estimates for tropical moist forests of the Brazilian Amazon. *Interciencia*, 17: 8 - 18.
- Buddenbaum, H., Schlerf, M. and Hill, J., 2005. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *International Journal of Remote Sensing*, 26(24): 5453-5465.
- Burrough, P.A., 1983. Multiscale sources of spatial variation in soil. II. A non- Brownian fractal model and its application in soil survey. *Journal of Soil Science*, 34(3): 599-620.
- Butler, R.A., 2006. Tropical Rainforest of the World. Mongabay.com / A Place Out of Time: Tropical Rainforests and the Perils They Face. [online] <http://rainforests.mongabay.com/0101.htm> (Accessed on: March 20 2008).
- Canty, M.J., 2009. Boosting a fast neural network for supervised land cover classification. *Computers and Geosciences*, 35(6): 1280-1295.

- Canty, M.J., Nielsen, A.A. and Schmidt, M., 2004. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sensing of Environment*, 91: 441 - 451.
- Carpenter, G.A., Gजा, M.N., Gopal, S. and Woodcock, C.E., 1997. Art neural networks for remote sensing: vegetation classification from landsat tm and terrain Data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(2): 308-325.
- Carr, J., 1995. *Numerical Analysis for the Geological Sciences*. Prentice - Hall, Inc., New Jersey.
- Chang, C.C. and Lin, C.J., 2001. LIBSVM: a library for support vector machines Software. [online] <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (Accessed on: September 19, 2007).
- Chaudhuri, B.B. and Sarkar, N., 1995. Texture Segmentation Using Fractal Dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(72-77).
- Chavez Jr., P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24: 459-479.
- Chavez Jr., P.S., 1996. Image-based atmospheric corrections - Revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62(9): 1025-1036.
- Chica-Olmo, M. and Abarca-Hernandez, F., 2000. Computing geostatistical image texture for remotely sensed data classification. *Computers and Geosciences*, 26(4): 373-383.
- Cloude, S. et al., 2008. Forest structure estimation using space borne polarimetric radar: An ALOS PALSAR case study, European Space Agency, (Special Publication) ESA SP. 2008 Dragon Symposium - Dragon 1 Programme Final Results 2004-2007, Beijing.
- Cohen, J. and Cohen, P., 1983. *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ.
- Collins, N.M., Sayer, J.A. and Whitmore, T.C. (Editors), 1991. *Conservation Atlas of Tropical Forests: Asia and the Pacific*. Simon & Schuster / IUCN, New York, 256 pp.
- Cooley, T. et al., 2002. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its applications and validation, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2002)*, Toronto, Ont., pp. 1414-1418.
- De Gier, A., 2003. A new approach to woody biomass assessment in woodlands and shrublands. In: P. Roy (Editor), *Geoinformatics for Tropical Ecosystems*, India, pp. 161-198.
- Demuth, H., Beale, M. and Hagan, M., 2006. *Neural Network Toolbox 5 User's Guide*. The MathWorks, Inc., Apple Hill Drive Natick, MA.

- Deutsch, C.V. and Journel, A.G., 1998. GSLIB: geostatistical software library and user's guide. Second edition. GSLIB: geostatistical software library and user's guide. Second edition. Oxford University Press, Oxford, New York.
- Drake, J.B. et al., 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecology and Biogeography*, 12: 147-159.
- Epting, J., Verbyla, D.L. and Sorbel, B., 2005. Evaluation of Remotely Sensed Indices for Assessing Burn Severity in Interior Alaska using Landsar TM and ETM+. *Remote Sensing of Environment*, 96(3-4): 328-339.
- European Space Agency (ESA), 2007. Tutorial of PolSarpro (Development of a Polarimetric SAR Image Analysis Tool, ESA-ESRIN contract no. 17863/03/I-LG. In: E. Pottier and L.F. Famil (Editors).
- FAO, 2003. State of the world's forests. [online] <http://www.fao.org/forestry/site/10168/en> (Accessed on: July 12 2006).
- FAO, 2004. Global forest resources assessment update 2005, Terms and definition. [online] <http://www.fao.org/docrep/007/ae156e/ae156e00.HTM> (Accessed on: November 17 2007).
- FAO, 2006. Global Forest Resources Assessment 2005 : Progress towards sustainable forest management, Rome.
- Fauzi, A., 2001. Remote Sensing for detecting tropical logged over forest: a case study in Labanan consession, East Kalimantan. Indonesia. MSc. Thesis, ITC The Netherlands, Enschede.
- Fearnside, P.M., 1997. Greenhouse gases from deforestation in Brazilian Amazonia: Net committed emissions. *Climatic Change*, 35(3): 321-360.
- Foody, G.M., 2003. Remote sensing of tropical forest environments: towards the monitoring of environmental resources for sustainable development. *International Journal of Remote Sensing*, 24(20): 4035-4046.
- Foody, G.M., Boyd, D.S. and Cutler, M.E.J., 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment*, 85(4): 463-474.
- Foody, G.M. and Cox, D.P., 1994. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, 15(3): 619-631.
- Foody, G.M. and Curran, P.J., 1994. Estimation of tropical forest extent and regenerative stage using remotely sensed data. *Journal of Biogeography*, 21: 223 - 244.

- Foody, G.M. et al., 2001. Mapping the biomass of Bornean tropical rain forest from remotely sensed data. *Global Ecology and Biogeography*, 10(4): 379-387.
- Foody, G.M. and Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103(2): 179-189.
- Fransson, J.E.S. and Israelsson, H., 1999. Estimation of stem volume in boreal forests using ERS-1 C- and JERS-1 L-band SAR data. *International Journal of Remote Sensing*, 20(1): 123-137.
- FWI/GFW, 2002. The State of the Forest: Indonesia. , Forest Watch Indonesia, and Washington DC: Global Forest Watch., Bogor, Indonesia.
- Gahegan, M., German, G. and West, G., 1999. Improving neural network performance on the classification of complex geographic datasets. *Journal of Geographical Systems*, 1(1): 3-22.
- Gillespie, T.W., Brock, J. and Wright, C.W., 2004. Prospects for quantifying structure, floristic composition and species richness of tropical forests. *International Journal of Remote Sensing*, 25(4): 707-715.
- Gloaguen, R., Marpu, P. and Niemeyer, I., 2007. Automatic Extraction of Faults and Fractal Analysis from Remote Sensing Data. *Nonlinear Processes Geophysics*, 14: 131-138.
- Hagan, M.T. and Menhaj, M.B., 1994. Training feedforward networks with the Marquardt algorithm. *Neural Networks, IEEE Transactions on*, 5(6): 989-993.
- Hahn, C., Wijaya, A. and Gloaguen, R., 2007. Application of Support Vector Machine for Complex Land Cover Classification using Aster and Landsat Data, *Proceedings of Gemeinsame Jahrestagung der SGPBF, DGPF und OVG, Publications of DGPF, Muttentz/Basel, Switzerland*, pp. 149-154.
- Hajnsek, I., Kugler, F., Lee, S.K. and Papathanassiou, K.P., 2009. Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2): 481-493.
- Hajnsek, I. et al., 2005. INDREX II - Indonesian airborne radar experiment campaign over tropical forest in L- and P-band: First results, *International Geoscience and Remote Sensing Symposium (IGARSS). 2005 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2005, Seoul*, pp. 4335-4338.
- Hame, T., Salli, A., Andersson, K. and Lohi, A., 1997. A new methodology for the estimation of biomass of conifer-dominated boreal forest using NOAA AVHRR data. *International Journal of Remote Sensing*, 18(15): 3211-3243.
- Haralick, R.M., Shanmugam, K. and Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, smc 3(6): 610-621.

- He, L.-M., Kong, F.-S. and Shen, Z.-Q., 2005. Multiclass SVM based land cover classification with multisource data, 2005 International Conference on Machine Learning and Cybernetics, ICMLC 2005. International Conference on Machine Learning and Cybernetics, ICMLC 2005, Guangzhou, pp. 3541-3545.
- Hecker, J.H., 2005. Promoting Environmental Security and Poverty Alleviation in the Peat Swamps of Central Kalimantan, Indonesia Institute for Environmental Security, The Hague, The Netherlands, The Netherlands.
- Hoekman, D.H., 2007. Satellite radar observation of tropical peat swamp forest as a tool for hydrological modelling and environmental protection. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 17(3): 265-275.
- Houghton, R.A., Lawrence, K.T., Hackler, J.L. and Brown, S., 2001. The spatial distribution of forest biomass in the Brazilian Amazon: A comparison of estimates. *Global Change Biology*, 7(7): 731-746.
- Hsu, C.-W. and Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2): 415-425.
- Huang, C., Davis, L.S. and Townshend, J.R.G., 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4): 725-749.
- Huete, A.R., 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25: 295-309.
- Huete, A.R., Liu, H., Batchily, K. and Leeuwen, W.v., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment* 59(3): 440-451.
- Hyypä, J. et al., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, 128(1-2): 109-120.
- Imhoff, M.L., 1995. Radar backscatter and biomass saturation: ramifications for global biomass inventory. *IEEE Transactions on Geoscience and Remote Sensing*, 33(2): 511-518.
- Isola, M. and Cloude, S.R., 2001. Forest height mapping using space-borne polarimetric SAR interferometry, *International Geoscience and Remote Sensing Symposium (IGARSS)*. 2001 International Geoscience and Remote Sensing Symposium (IGARRS 2001), Sydney, NSW, pp. 1095-1097.
- Jaenisch, H.M., Barton, P.E. and Carruth, R.T., 1993. Determining the fractal dimension of scenes and digital signals using ROSETA and other novel approaches. In: I. Kadar and V. Libby (Editors), *Proceedings of SPIE Signal Processing, Sensor Fusion, and Target Recognition II*, pp. 298-315.

- Jakomulska, A. and Clarke, K.C., 2001. Variogram-Derived Measures of Textural Image Classification. In: P. Monesties (Editor), *geoENV III - Geostatistics for Environment Applications*. Kluwer Academic Publisher, Avignon, pp. 345-355.
- Jensen, J.R., 1996. *Introductory Digital Image Processing : A remote Sensing Perspective*. Prentice Hall.
- Journel, A.G. and Huijbregts, C.J., 1981. *Mining Geostatistics*. Academic Press.
- Kanellopoulos, I. and Wilkinson, G.G., 1997. Strategies and best practice for neural network image classification. *International Journal of Remote Sensing*, 18(4): 711-725.
- Kaufman, Y.J. and Tanre, D., 1996. Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: from AVHRR to EOS-MODIS. *Remote Sensing of Environment*, 55: 65-79.
- Kayitakire, F., Hamel, C. and Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, 102(3-4): 390-401.
- Ketterings, Q.M., Coe, R., Van Noordwijk, M., Ambagau, Y. and Palm, C.A., 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *Forest Ecology and Management*, 146(1-3): 199-209.
- Kimes, D.S., Nelson, R.F., Manry, M.T. and Fung, A.K., 1998. Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements. *International Journal of Remote Sensing*, 19(14): 2639-2662.
- Kingsley, S. and Quegan, S., 1992. *Understanding radar systems*. SciTech Publishing, 384 pp.
- Kuplich, T.M., 2006. Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. *Forest Ecology and Management*, 234(1-3): 1-9.
- Kuplich, T.M., Curran, P.J. and Atkinson, P.M., 2005. Relating SAR image texture to the biomass of regenerating tropical forests. *International Journal of Remote Sensing*, 26(21): 4829-4854.
- Landsat Project Science Office, 1998. *Landsat 7 Science Data Users Handbook*. [online] <http://landsathandbook.gsfc.nasa.gov/handbook/handbook.htmls/chapter11/chapter11.html> (Accessed on: November 7 2007).
- Lefsky, M.A. et al., 2002. Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography*, 11(5): 393-399.

- Li, M., Tan, Y., Pan, J. and Peng, S., 2008. Modeling forest aboveground biomass by combining spectrum, textures and topographic features. *Frontiers of Forestry in China*, 3(1): 10-15.
- Li, Y., Moran, E.F., Brondizio, E.S., Mausel, P. and Wu, Y., 1994. Discrimination between advanced secondary succession and mature moist forest near Altamira, Brazil using Landsat TM data, In: *Proceedings of the American Society for Photogrammetry and Remote Sensing and Annual Meeting of ASPRS*, Reno, USA (November 1994).
- Liesenberger, V., Boehm, H.-D.V. and Gloaguen, R., 2009. The Contribution of Chris/Proba Data for Tropical Peat Swamp Landscape Discrimination Purposes, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Cape Town, South Africa.
- Lillesand, T.M. and Kiefer, R.W., 1994. *Remote Sensing and Image Interpretation*. John Wiley & Sons, Inc.
- Liu, J., Shao, G., Zhu, H. and Liu, S., 2005. A neural network approach for enhancing information extraction from multispectral image data. *Canadian Journal of Remote Sensing*, 31(6): 432-438.
- Lloyd, C.D., Berberoglu, S., Curran, P.J. and Atkinson, P.M., 2004. A comparison of texture measures for the per-field classification of Mediterranean land cover. *International Journal of Remote Sensing*, 25(19): 3943-3965.
- Losi, C.J., Siccama, T.G., Condit, R. and Morales, J.E., 2003. Analysis of alternative methods for estimating carbon stock in young tropical plantations. *Forest Ecology and Management*, 184(1-3): 355-368.
- Lu, D., 2005. Aboveground biomass estimation using Landsat TM data in Brazilian Amazon. *International Journal of Remote Sensing*, 26(12): 2509-2525.
- Lu, D. and Batistella, M., 2005. Exploring TM Image Texture and its Relationships with Biomass Estimation in Rondônia, Brazilian Amazon. *Acta Amazonica*, 35(2): 249 - 257.
- Lu, D., Mausel, P., Brondizio, E. and Moran, E., 2003. Classification of successional forest stages in the Brazilian Amazon basin. *Forest Ecology and Management*, 181(3): 301-312.
- Lu, D., Mausel, P., Brondizio, E. and Moran, E., 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. *Forest Ecology and Management*, 198(1-3): 149-167.
- Lu, D.S., 2006. The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing*, 27(7): 1297-1328.
- Lucas, R.M., Honzak, M., do Amaral, I., Curran, P.J. and Foody, G.M., 2002. Forest regeneration on abandoned clearances in central Amazonia. *International Journal of Remote Sensing*, 23: 965-988.

- Lucas, R.M., Honzak, M., Foody, G.M., Curran, P.J. and Corves, C., 1993. Characterizing tropical secondary forests using multi-temporal Landsat sensor imagery. *International Journal of Remote Sensing*, 14(16): 3061 - 3067.
- Luckman, A., Baker, J., Kuplich, T.M., Corina da Costa, F.Y. and Alejandro, C.F., 1997. A study of the relationship between radar backscatter and regenerating tropical forest biomass for spaceborne SAR instruments. *Remote Sensing of Environment*, 60(1): 1-13.
- Luckman, A., Baker, J., Kuplich, T.M., Yanasse, C.d.C.F. and Frery, A.C., 1996. A study of the relationship between radar backscatter and regenerating tropical forest biomass for spaceborne SAR instruments. *Remote Sensing of Environment*, 60(1): 1-13.
- Mantel, S., 1998. Soil and Terrain of the Labanan Area: Development of an environmental framework for the Berau Forest Management Project. Berau Forest Management Project, Berau, 142 pp pp.
- Marpu, P.R., Wijaya, A. and Gloaguen, R., 2008. Soft Classification and Assessment of Kalman Filter Neural Network for Complex Landcover of Tropical Rainforests, *Geoscience and Remote Sensing Symposium*, 2008. IGARSS 2008. IEEE International, pp. V - 57-V - 60.
- Mather, P.M., 2004. *Computer Processing of Remotely-Sensed Data : An Introduction*. John Wiley & Sons, Ltd, Chichester, West Sussex.
- Mausel, P., Wu, Y., Li, Y., Moran, E.F. and Brondizio, E.S., 1993. Spectral identification of succession stages following deforestation in the Amazon. *Geocarto Int.*, 8: 61-72.
- McNaught, A.D. and A.Wilkinson, 1997. *IUPAC. Compendium of Chemical Terminology. (the "Gold Book")*. Blackwell Scientific Publications, Oxford.
- Mette, T., Hajnsek, I. and Papathanassiou, K., 2003. Height-Biomass Allometry in Temperate Forests: Performance accuracy of height-biomass allometry, *International Geoscience and Remote Sensing Symposium (IGARSS)*. 2003 IGARSS: Learning From Earth's Shapes and Colours, Toulouse, pp. 1942-1944.
- Minnaert, M., 1941. The reciprocity principle in lunar photometry. *Astrophysical Journal*, 93: 403-410.
- Miranda, F.P. and Carr, J.R., 1994. Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of the Guiana Shield, northwestern Brazil. *Remote Sensing Reviews*, 10(1-3): 155-168.
- Miranda, F.P., Fonseca, L.E.N. and Carr, J.R., 1998. Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. *International Journal of Remote Sensing*, 19(3): 549-556.
- Miranda, F.P., Fonseca, L.E.N., Carr, J.R. and Taranik, J.V., 1996. Analysis of JERS-1 (Fuyo-1) SAR data for vegetation discrimination in northwestern Brazil using the

- semivariogram textural classifier (STC). *International Journal of Remote Sensing*, 17(17): 3523-3529.
- Miranda, F.P., Macdonald, J.A. and Carr, J.R., 1992. Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo. *International Journal of Remote Sensing*, 13(12): 2349-2354.
- MoF, 2002. Info Kehutanan. Indonesian Ministry of Forestry (MoF), Jakarta.
- MoF, 2006. Forestry Statistical Book. Ministry Forestry of Indonesia, Jakarta, 168 pp.
- Moran, E.F. and Brondizio, E.S., 1998. Land use change after deforestation in Amazonia. In: D. Liverman, E.F. Moran, R.R. Rindfuss and P.C. Stern (Editors), *People and Pixels : Linking Remote Sensing and Social Science*. National Academy Press, Washington, DC., pp. 94 - 120.
- Moran, E.F., Brondizio, E.S., Mausel, P. and Wu, Y., 1994. Integrating Amazonian vegetation, land use, and satellite data. *BioScience*, 44: 329-338.
- Moran, E.F. et al., 2000. Effects of soil fertility and land-use on forest succession in Amazonia. *Forest Ecology & Management*, 139: 93 - 108.
- Muhamad, N.Z. and Rieley, J.O., 2002. Management of tropical peatlands in Indonesia: mega reclamation project in Central Kalimantan. *Peatlands for People: Natural Resource Functions and Sustainable Management*. , Proceedings of the International Symposium on Tropical Peatland. BPPT and Indonesian peat Association, Jakarta, pp. 155-167.
- Muukkonen, P. and Heiskanen, J., 2005. Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data *Remote Sensing of Environment*, 99: 434 - 447.
- Nielsen, A.A., Conradsen, K. and Simpson, J.J., 1998. Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, 64: 1 - 19.
- Nykvist, N., 1996. Regrowth of secondary vegetation after the 'Borneo Fire' of 1982-1983. *Journal of Tropical Ecology*, 12: 307-312.
- Page, S.E. et al., 2002. The amount of carbon released from peat and forest fires in Indonesia during 1997. *Nature* 420: 61-65.
- Pal, M. and Mather, P.M., 2005. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26(5): 1007-1011.
- Parresol, R., 1999. Assessing Tree and Stand Biomass: A Review with Examples and Critical Comparisons. *Forest Science*, 45: 573-593.

- Pinty, B. and Verstraete, M.M., 1991. GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetation*, 101: 15-20.
- Qi, J., Kerr, Y. and Chehbouni, A., 1994. External Factor Consideration in Vegetation Index Development, *Proceeding of Physical Measurements and Signatures in Remote Sensing*, ISPRS, pp. 723-730.
- Rahman, M.M., Csaplovics, E. and Koch, B., 2005. An efficient regression strategy for extracting forest biomass information from satellite sensor data. *International Journal of Remote Sensing*, 26(7): 1511-1519.
- Ramankutty, N. et al., 2007. Global Land-Cover Change: Recent Progress, Remaining Challenges In: E.F. Lambin and H. Geist (Editors), *Land-Use and Land-Cover Change Local Processes and Global Impacts : Global Change - The IGBP Series*. Springer Berlin Heidelberg, pp. 9 - 39.
- Ranson, K.J. et al., 1997. Mapping of boreal forest biomass from spaceborne synthetic aperture radar. *Journal of Geophysical Research D: Atmospheres*, 102(24): 29599-29610.
- Rauste, Y., 2005. Multi-temporal JERS SAR data in boreal forest biomass mapping. *Remote Sensing of Environment*, 97(2): 263-275.
- Ravindranath, N.H. and Ostwald, M., 2007. Methods for Below-Ground Biomass, Carbon Inventory Methods Handbook for Greenhouse Gas Inventory, Carbon Mitigation and Roundwood Production Projects. *Advances in Global Change Research*. Springer Netherlands, pp. 149-156.
- Rees, W.G., 2001. *Physical Principles of Remote Sensing*. Cambridge University Press.
- Riaño, D., Chuvieco, E., Salas, J. and Aguado, I., 2003. Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing*, 41(5): 1056 - 1061.
- Richards, J.A., 1993. *Remote Sensing Digital Image Analysis - An Introduction*. Springer-Verlag, New York.
- Richter, R., 1996. Atmospheric correction of satellite data with haze removal including a haze/clear transition region. *Computers and Geosciences*, 22(6): 675-681.
- Rieley, J.O. and Page, S.E. (Editors), 2005. *Wise use of Tropical Peatlands: Focus on Southeast Asia*. Synthesis of results and conclusions of the UK Darwin Initiative and the EU INCO EUTROP, STRAPEAT and RESTORPEAT Partnerships together with proposals for implementing wise use of tropical peatlands. ALTERRA - Wageningen University and Research Centre and the EU INCO - STRAPEAT and RESTORPEAT Partnerships, Wageningen, The Netherlands, 266pp pp.

- Rignot, E., Salas, W.A. and Skole, D.L., 1997. Mapping deforestation and secondary growth in Rondonia, Brazil, using imaging radar and thematic mapper data. *Remote Sensing of Environment*, 59(2): 167-179.
- Rogers, G.W., Solka, J.L. and Priebe, C.E., 1995. A PDP Approach to Localized Fractal Dimension Computation with Segmentation Boundaries. *SIMULATION*, 65(1): 26-36.
- Rosenqvist, A., Shimada, M. and Milne, A.K., 2008. The ALOS Kyoto & carbon initiative, International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3614-3617.
- Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. , Third ERTS Symposium, NASA SP-351 I, pp. 309-317.
- Roy, P.S. and Ravan, S.A., 1996. Biomass estimation using satellite remote sensing data - An investigation on possible approaches for natural forest. *Journal of Biosciences*, 21(4): 535-561.
- Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. Learning internal Representations by Error Propagation. In: D.E. Rumelhart, J.L. McClelland and P.R. Group (Editors), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press, Cambridge, pp. 318-362.
- Saatchi, S.S., Soares, J.V. and Alves, D.S., 1997. Mapping deforestation and land use in amazon rainforest by using SIR-C imagery. *Remote Sensing of Environment*, 59(2): 191-202.
- Sader, S.A., Waide, R.B. and Lawrence, W.T., 1989. Tropical Forest Biomass and Successional age Class Relationship to a Vegetation Index Derived from Landsat TM Data. *Remote Sensing of Environment*, 28: 143-156.
- Salas, W.A., Ducey, M.J., Rignot, E. and Skole, D., 2002. Assessment of JERS-1 SAR for monitoring secondary vegetation in Amazonia: I. Spatial and temporal variability in backscatter across a chrono-sequence of secondary vegetation stands in Rondonia. *International Journal of Remote Sensing*, 23(7): 1357-1379.
- Sales, M.H., Souza Jr., C.M., Kyriakidis, P.C., Roberts, D.A. and Vidal, E., 2007. Improving spatial distribution estimation of forest biomass with geostatistics: A case study for Rondonia, Brazil. *Ecological Modelling*, 205(1-2): 221-230.
- Samalca, I., 2007. Estimation of Forest Biomass and its Error: A case in Kalimantan, Indonesia. Unpublished MSc. Thesis Thesis, ITC the Netherlands, Enschede, 84 pp.
- Sari, A.P., Maulidya, M., Butarbutar, R.N., Sari, R.E. and Rusmantoro, W., 2007. Working Paper: Indonesia and Climate Change 'Current Status and Policies', DFID and World Bank.
- Scholkopf, B. and Smola, A., 2001. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press, Cambridge, MA, USA.

- Schroeder, T.A., Cohen, W.B., Song, C., Canty, M.J. and Yang, Z., 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. *Remote Sensing of Environment*, 103(1): 16-26.
- Sist, P. and Nguyen-Thé, N., 2002. Logging damage and the subsequent dynamics of a dipterocarp forest in East Kalimantan (1990-1996). *Forest Ecology and Management*, 165(1-3): 85-103.
- Song, C., Schroeder, T.A. and Cohen, W.B., 2007. Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sensing of Environment*, 106: 228 - 237.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P. and Macomber, S.A., 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, 75(2): 230-244.
- Steininger, M., 2000. Satellite estimation of tropical secondary forest above-ground biomass: data from Brazil and Bolivia. *International Journal of Remote Sensing*, 21(6 & 7): 1139-1157.
- Steininger, M.K., 1996. Tropical secondary forest regrowth in the Amazon: age, area and change estimation with Thematic Mapper data. *International Journal of Remote Sensing*, 17: 9-27.
- Stibig, H.J. and Malingreau, J.P., 2003. Forest Cover of Insular Southeast Asia Mapped from Recent Satellite Images of Coarse Spatial Resolution. *A Journal of the Human Environment*, 32(7): 469-475.
- Susilo, H.D., 1997. The Tanjung Puting National Park and Biosphere Reserve Indonesia (Working Papers no. 22, 1997), Division of Ecological sciences, UNESCO, Paris, France.
- Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S., 2001. Multiple-class land-cover mapping at the sub-pixel scale using a Hopfield neural network. *International Journal of Applied Earth Observation and Geoinformation*, 3(2): 184-190.
- Teillet, P.M., Guindon, B. and Goodenough, D.G., 1982. On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing*, 8(2): 84-106.
- Thenkabail, P.S. et al., 2004. Biomass estimations and carbon stock calculations in the oil palm plantations of African derived savannas using IKONOS data. *International Journal of Remote Sensing*, 25(23): 5447-5472.
- Tso, B. and Mather, P.M., 2001. *Classification Methods for Remotely Sensed Data*. Taylor and Francis, London.
- Tucker, J.M., Brondizio, E.S. and Moran, E.F., 1998. Rates of forest regrowth in Eastern Amazonia: a comparison of Altamira and Bragantina regions, Pará State, Brazil. *Interciencia* 23: 64-73.

- Uhl, C., Buschbacher, R. and Serrao, E.A.S., 1988. Abandoned Pastures in Eastern Amazonia. I. Patterns of Plant Succession. *Journal of Ecology*, 76(663 - 681).
- UNFCCC, 1992. United Nations Framework Convention on Climate Change. FCCC/Informal/84 GE.05-62220 (E) 200705. [online] <http://unfccc.int/resource/docs/convkp/conveng.pdf> (Accessed on: 1 September 2009).
- van Beukering, P.J.H., Schaafsma, M., Davies, O. and Oskolokaite, I., 2008. The economic value of peatland resources within the Central Kalimantan Peatland Project in Indonesia: Perceptions of local communities. 29 June 2008, Central Kalimantan Peatlands Project. Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands.
- Vapnik, V.N., 1999. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5): 988-999.
- Vapnik, V.N., 2000. The nature of statistical learning theory. Springer Verlag.
- Venkataraman, G. et al., 2004. Fusion of optical and microwave remote sensing data for snow cover mapping. *Igarss 2004: Ieee International Geoscience and Remote Sensing Symposium Proceedings, Vols 1-7 - Science for Society: Exploring and Managing a Changing Planet*: 2554-2557.
- Vieira, I.C.G. et al., 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sensing of Environment*, 87(4): 470-481.
- Wahyunto, Ritung, S. and Subagjo, H., 2004. Peta luas sebaran lahan gambut dan kandungan karbon di Pulau Kalimantan 2000 - 2002 (Maps of area of peatland distribution and carbon content in Kalimantan 2000 - 2002) (In Indonesian), Wetlands International - Indonesia Programme & Wildlife Habitat Canada (WHC), Bogor.
- Wetlands International, 2004. Peta luas sebaran lahan gambut dan kandungan karbon di Pulau Kalimantan 2000 - 2002 (Maps of area of peatland distribution and carbon content in Kalimantan 2000 - 2002) (In Indonesian), Wetlands International - Indonesia Programme, Bogor.
- Wetlands International, 2007. Assessment on Peatlands, Biodiversity and Climate Change.
- Widayat, A.S., 2005. Accessing research data to support research activities for forest management: a case study of MONCER project information system design, ITC The Netherlands, Enschede.
- Wielgaard, N. and Hoekman, D., 2009. Update radar mapping results: Indonesia and Guyana Second GEO Forest Monitoring Symposium: from research to operations - Session 8, Chiang Mai, Thailand.

- Wijaya, A., 2006. Comparison of soft classification techniques for forest cover mapping. *Journal of Spatial Science*, 51(2): 7-18.
- Wijaya, A. and Gloaguen, R., 2007a. Application of Support Vector Machine for Forest Cover Mapping using Multi-Source Remotely Sensed Data (in CD ROM), IGARSS 'Sensing and Understanding our Planet', Barcelona, Spain.
- Wijaya, A. and Gloaguen, R., 2007b. Comparison of multisource data support vector Machine classification for mapping of forest cover, *Geoscience and Remote Sensing Symposium*, 2007. IGARSS 2007. IEEE International, pp. 1275-1278.
- Wijaya, A. and Gloaguen, R., 2008. Spatial estimation of stand volume using integration of field measurement and remote sensing data of tropical rainforest in central Indonesia, In *Proceedings Of Gemeinsame Jahrestagung Der SGPBF, DGPF Und OVG. DGPF, Oldenburg*, pp. 189-198.
- Wijaya, A., Gloaguen, R. and Heilmeyer, H., 2008a. Integration of field measurement and remotely sensed data for estimation of stand volume and above ground biomass of tropical rainforest in Indonesia, In *LIFO International Conference 'Linking Forest Inventory And Optimization'*, Freising.
- Wijaya, A., Marpu, P.R. and Gloaguen, R., 2007. Geostatistical Texture Classification of Tropical Rainforest in Indonesia, *ISPRS International Symposium on Spatial Data Quality*, ITC Enschede, The Netherlands (online: <http://www.itc.nl/ISSDQ2007/proceedings>).
- Wijaya, A., Marpu, P.R. and Gloaguen, R., 2008b. Geostatistical texture classification of tropical rainforest in Indonesia. In: J.S. Alfred Stein, and Wietske Bijker (Editor), *Quality Aspect in Spatial Data Mining*. CRC Press Inc., pp. 199-210.
- Wijaya, A., Sharifi, M.A. and Tolpekin, V., 2005. Integration of Remotely Sensed Data and Experts Knowledge to Detect Illegal Logging in the Production Forest of East Kalimantan, Indonesia (Abstract), *Proceeding on the IUFRO Session 135 "Detecting, Monitoring and Modelling Deforestation and Forest Degradation Using Remote Sensing and GIS"*, XXII IUFRO World Congress: Forest in the Balance: Linking Tradition & Technology, 8-13 August 2005, Brisbane, Australia, pp. 187.
- Wolter, P.T., Mladenoff, D.J., Host, G.E. and Crow, T.R., 1995. Improved Forest Classification in the Northern Lake States using Multi-Temporal Landsat Imagery. *Photogrammetric Engineering & Remote Sensing*, 61(9): 1129-1143.
- Woodcock, C.E., Strahler, A.H. and Jupp, D.L.B., 1988. The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, 25(3): 349-379.
- Wu, M.-c. and Chen, K.-S., 2007. Fusion of SPOT and SAR Images for Land Cover Classification. *Journal of Photogrammetry and Remote Sensing*, 12(1): 59-72.

- Wulder, M.A., 1998. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Progress in Physical Geography*, 22(4): 449-476.
- Yool, S.R., 1998. Land cover classification in rugged areas using simulated moderate resolution remote sensor data and an artificial neural network. *International Journal of Remote Sensing*, 19(1): 85-96.
- Zhang, C., Franklin, S.E. and Wulder, M.A., 2004. Geostatistical and texture analysis of airborne-acquired images used in forest classification. *International Journal of Remote Sensing*, 25(4): 859-865.

Appendix 1 Curriculum Vitae

ARIEF WIJAYA

REMOTE SENSING FOR
FORESTRY & ENVIRONMENTAL MODELING



EDUCATION

- 2006 – Present *Doctoral research*
Remote Sensing Group, Technische Universität Bergakademie,
Freiberg, Germany
Research Topic : Non-destructive approach for biomass
assessment using remote sensing and non-parametric
methods
Degree Awarded : Doctor rerum naturalium (Dr rer nat)
Date received : April 30, 2010
- 2003 – 2005 *Masters study*
International Institute for Geo-Information Science and
Earth Observation (ITC), Enschede, The Netherlands
Department of Natural Resource Management
Degree Awarded: Master of Science (M.Sc)
Date received : March 10, 2005
- 1999 – 2001 *Postgraduate study*
Master of Management Program
Major: Marketing
Gadjah Mada University, Yogyakarta, Indonesia
Degree Awarded: Master of Management (MM)
Date received : September 25, 2001
- 1993 – 1998 *Undergraduate Study*
Agricultural Technology Faculty
Gadjah Mada University, Yogyakarta, Indonesia
*Degree Awarded: Sarjana (bachelor) in Agro-Industrial
Technology Department (STP)*
Date Received : November 19, 1998

WORK EXPERIENCE

- 2003 – 2005 International Institute for Geoinformation Science & Earth
Observation (ITC), Enschede, The Netherlands
*M.Sc. Student and Researcher for Monitoring & Certification of Tropical rainforest
Project in Indonesia*
- Extensive learning of remotely sensed data (optical and microwave),
techniques, and applications.

- Extensive learning of Geostatistics applying Kriging methods for the prediction of stand volume of Indonesian tropical forest
- The MSc thesis topic was the application of remotely sensed data and methods for detecting illegal logging in the tropics using Landsat data. ERDAS and ENVI software were used for image classification applying fuzzy and neural network methods. The results were combined with expert knowledge and relevant GIS maps for detecting the illegally logged forests.

2000 – 2003 Department of Industrial Management and Engineering, Faculty of Industrial Engineering, Indonesia Islamic University, Yogyakarta, Indonesia

Visiting Lecturer

Teaching Subjects :

- Marketing (e.g., principles of marketing, marketing strategy, consumer behavior, marketing distribution system, and new product development)
- Decision Making Analysis (e.g., decision making modeling, Bayesian theory, decision tree, analytical hierarchy process, and probabilistic decision making)

1998 – 2003 Department of Agro industrial Technology, Faculty of Agricultural Technology Gadjah Mada University, Yogyakarta, Indonesia

Lecturer and Researcher

Supporting Associate Professor at the Industrial System and Management Laboratory, as well as at the Simulation System and Analysis Laboratory. Very comprehensive learning and experience of fuzzy and neural network techniques for agricultural system and model development.

Teaching Subjects :

- Operations Research (e.g., linear programming, simplex method, dynamic programming, nonlinear programming, probabilistic theory, network models, and project management models)
- Computer Application and software (dealing with Microstat & MathCad)
- Decision Making Analysis (e.g., decision making modeling, Bayesian theory, decision tree, analytical hierarchy process, and probabilistic decision making)

PROJECT EXPERIENCE

Sept 2004 – March 2005 Monitoring and Certification of Tropical Forest in Indonesia (MONCER) Project. Task: To identify illegal forest harvesting using integration of remote sensing and GIS data.

Aug 2008 – Dec 2008 Project: Tropical forest deforestation and sustainability of Sumatran Tiger : Individual based modeling approach. Task: Spatial data and GIS layers preparation, land cover map classification, extraction of river drainage from digital elevation model.

Sept 2009 – Dec 2009 Project: Developing carbon-incentive mechanisms for forest resource management toward the emission reduction in Kampar Peninsular peatlands, Indonesia: Establishing baseline activities. Task: remote sensing specialist using radar and optical data, to establish remote sensing – forest biomass allometric model, and to compose an article for publication in international peer-reviewed journal

SCHOLARSHIPS AND GRANTS

2003	STUNED Scholarship, Netherlands Education Center, Master study at ITC the Netherlands (2003 – 2005)
2006	DAAD Scholarship, German Academic Exchange Service, PhD Study at TU Bergakademie Freiberg, Germany (2006 – 2009)
March 2007	Research Grants, PhD Program TU Bergakademie Freiberg, for conducting field work in Kalimantan, Indonesia
July 2007	Travel Grants, PhD Program TU Bergakademie Freiberg, Oral Talk at IGARSS 2008, Barcelona, Spain.
March 2008	Research Grants, PhD Program TU Bergakademie Freiberg, for conducting field visit in Kalimantan, Indonesia
July 2008	IEEE Travel Grants, Oral Talk at IGARSS 2008, Boston, USA

PROFESSIONAL MEMBERSHIP

2006 – Present	<i>Student Member</i> , Institute of Electrical and Electronics Engineers (IEEE)
2006 – Present	<i>Member</i> , International Association of Mathematical Geoscience (IAMG)
2007 – Present	<i>Member</i> , IEEE Geoscience and Remote Sensing Society
2006 – Present	<i>Activity Board</i> , IAMG Student Chapter, Freiberg, Germany

PERSONAL DATA

Place/Date of Birth	: Jakarta/March 24, 1975
Nationality	: Indonesian

I certify that this statement has explained about my resume and it is based on the truth. I will be responsible for any information given.

Freiberg, 17 May 2010

Sincerely,



Arief Wijaya

Appendix 2 Publication List

Peer-Reviewed Journals

Wijaya, A., P.R. Marpu and R. Gloaguen. *Discrimination of Peatlands in Tropical Swamp Forests using Dual-Polarimetric SAR and Landsat ETM Data*. International Journal of Image and Data Fusion, Taylor and Francis Publisher (In Press).

Wijaya, A., V. Liesenberg and R. Gloaguen (2010). *Retrieval of forest attributes in complex successional forests of Central Indonesia: Modeling and estimation of bitemporal data*. Forest Ecology and Management. Vol.259, Page 2315 – 2326

Wijaya, A., S. Kusnadi, R. Gloaguen and H. Heilmeyer. (2010). *Improved Strategy for Estimating Stem Volume and Forest Biomass using Moderate Resolution Remote Sensing Data and GIS*. Journal of Forestry Research, Vol.21 No.1, Page 1-12

Wijaya, A. (2009). Evaluation of ALOS Palsar Mosaic Data for Estimating Stem Volume and Biomass: a Case Study from Tropical Rainforest of Central Indonesia. Jurnal Geografi, Vol. 2 No. 1, January 2009, University of Indonesia, Jakarta, page 14 – 21.

Wijaya, A. (2006). Comparison of Soft Classification Techniques for Forest Cover Mapping. Journal of Spatial Science. Volume 51, Issue 2, December 2006, Page 7-18

Book Chapter

Wijaya, A., P.R. Marpu and R. Gloaguen (2008). Geostatistical Texture Classification of Tropical Rainforest in Indonesia. *Quality in Spatial Data Mining*. Editor: Stein, A., Shi, S. and Bijker, W. CRC Book Series, page 199 – 210.

Conference Proceedings

Wijaya, A., P.R. Marpu, V. Liesenberg and R. Gloaguen. *Assessing Peatlands in Tropical Swamp Forest of Central Indonesia with Dual-Polarimetric TerraSAR-X. data* (2010). European Space Agency (ESA) Living Planet Symposium, 28 June – 2 July 2010, Bergen, Norway.

Bayarbaatar, D., V. Liesenberg, **A. Wijaya** and R. Gloaguen (2010). *Evaluating the Angular and Spectral Properties of Some Land Cover Types in Huixian Site (China) using Chris/Proba Data*. European Space Agency (ESA) Living Planet Symposium, 28 June – 2 July 2010, Bergen, Norway.

Mahmood, S.A., V. Liesenberg, **A. Wijaya**, F. Shahzad and R. Gloaguen (2010). *Spectral Analysis of Hyperion, ALL, SPOT and ASTER data to characterize geological features in Northern Pakistan*. European Space Agency (ESA) Living Planet Symposium, 28 June – 2 July 2010, Bergen, Norway.

Mahmood, S.A., V. Liesenberg, **A. Wijaya**, F. Shahzad and R. Gloaguen (2010). *The Potential of Hyperion data to Discriminate Geological Features in the Eastern Hindukush Northern*

Pakistan. European Space Agency (ESA) Hyperspectral Workshop, March 17 – 19, 2010, ESRIN, Frascati, Italy.

Bouaziz, M., **A. Wijaya** and R. Gloaguen (2009). *Gully Erosion Mapping using Aster Data and Drainage Network Analysis in the Main Ethiopian Rift*. In the Proceeding of the International Geoscience and Remote Sensing Symposium (IGARSS), July 13 – 17, 2009, Cape Town, South Africa, page I-13 – I-16.

Wijaya, A. and R. Gloaguen (2009). *Fusion of ALOS Palsar and Landsat ETM data for land cover classification and biomass modeling using non-linear methods*. In the Proceeding of the International Geoscience and Remote Sensing Symposium (IGARSS), July 13 – 17, 2009, Cape Town, South Africa, page III-581 - III-584.

Wijaya, A. and R. Gloaguen (2009). *Multi Source Remotely Sensed Data for Land Cover Classification and Biomass Estimation using Non Linear Methods*. In the Proceeding of German Society for Photogrammetry, Remote Sensing and Geo-information (DGPF), March 24 – 26, 2009, Jena, Germany, page 207 – 216.

Wijaya, A. and R. Gloaguen (2008). *Using Combination of Remote Sensing and Field Observation Data for Mapping of Forest Stem Volume*. In the Proceeding of the Annual Scientific Meeting of Indonesian Remote Sensing Society 2008 (PIT-MAPIN), December 2008, Bandung, Indonesia, page 7 – 16.

Wijaya, A. and R. Gloaguen (2008). *Spatial Estimation of Stand Volume using Integration of Field Measurement and Remote Sensing Data of Tropical Rainforest in Central Indonesia*. In the Proceeding of the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF), April 23-25, 2008, Oldenburg, Germany, page 189 – 198.

Marpu, P.R., **A. Wijaya** and R. Gloaguen (2008). *Soft Classification and Assessment of Kalman Filter Neural Network for Complex Landcover of Tropical Rainforests*. In the Proceeding of the International Geoscience and Remote Sensing Symposium (IGARSS), July 2008, Boston, USA, page V-57 – V-60.

Wijaya, A., P.R. Marpu and R. Gloaguen (2007). *Geostatistical Texture Classification of Tropical Rainforest in Indonesia*. In the Proceeding of the 5th ISPRS International Symposium on Spatial Data Quality, June 2007, Enschede, The Netherlands, (online: <http://www.itc.nl/ISSDQ2007/proceedings>).

Hahn, C., **A. Wijaya** and R. Gloaguen (2007). *Application of Support Vector Machine for Complex Land Cover Classification using Aster and Landsat Data*. In the Proceeding of the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF), June 2007, in Muttentz/Basel, Switzerland, page 149-154.

Wijaya, A. and Richard Gloaguen (2007). *Application of Support Vector Machine for Forest Cover Mapping using Multi-Source Remotely Sensed Data*. In the Proceeding of the IGARSS 'Sensing and Understanding our Planet', July 2007, Barcelona, Spain, page 1275-1278.

Presentation in Conferences

Wirawan, B.A., N. Maulia and **A. Wijaya** (2009). *Fuzzy Logic and Cluster Shape Combination for Complex Land Cover Mapping of Aster Data*. In the Proceeding of the International Mathematical Geoscience (MatGeos): Geosciences – From Earth to Space, December 7 – 8, 2009, TU-Bergakademie, Freiberg, Germany

Imron, M.A., **A. Wijaya** and H. Suryatmojo (2009). *Geospatial Tools for Habitat Suitability Map of the Sumatran Tiger and its Prey*. In the Proceeding of the International Mathematical Geoscience (MatGeos): Geosciences – From Earth to Space, December 7 – 8, 2009, TU-Bergakademie, Freiberg, Germany

Babacan, K, M. Bouaziz, **A. Wijaya** and R. Gloaguen (2009). *Wavelet Analysis to Improve Gully Erosion Classification from Multispectral Data*. In the Proceeding of the International Mathematical Geoscience (MatGeos): Geosciences – From Earth to Space, December 7 – 8, 2009, TU-Bergakademie, Freiberg, Germany

Wijaya, A., R. Gloaguen, and H. Heilmeyer (2008). *Integration of Field Measurement and Remotely Sensed Data for Estimation of Stand Volume and Above Ground Biomass of Tropical Rainforest in Indonesia*. In the Proceeding of LIFO International Conference 'Linking Forest Inventory and Optimization', April 1-4, 2008, Freising, Germany.

Wijaya, A. and R. Gloaguen (2008). *Estimation of Forest Stand Properties through Remote Sensing Data Pruning and Levenberg-Marquardt Neural Network*. In the Proceeding of the International Mathematical Geoscience (MatGeos) Forum, June 11 – 12, 2008, TU-Bergakademie, Freiberg, Germany.

In Review

Wijaya, A., R. Gloaguen and H. Heilmeyer (Forthcoming). *Non-Destructive Approach Development using Levenberg-Marquardt Neural Network for Estimating Stand Volume and Above Ground Biomass in Secondary Tropical Forests of Indonesia*. Article is submitted to International Journal of Remote Sensing, Taylor and Francis Publisher.